

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**



GenBank version 9.1.4  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:15:05 : Search time 29.0579 Seconds  
(without alignments)  
350.688 Million cell updates/sec

Title: US-09-893-615-89

Sequence: 1 QIVLSQSPAILSPGKVT.....COOMSSNPPTFGCGTMLR 106

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum hit seq length: 0  
Maximum hit seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR\_73:\*

1: PIR1:  
2: PIR2:  
3: PIR3:  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	523	94.1	106	2	PL0082
2	492	88.5	107	2	A42848
3	480	86.3	107	2	PC4405
4	474	85.3	107	2	A30562
5	473	85.1	105	2	S26338
6	469	84.4	107	2	B30562
7	465.5	83.9	108	2	G30560
8	465.5	83.7	108	2	S38720
9	462	83.1	140	2	PL0013
10	459	82.6	106	2	PS0071
11	458	82.4	104	2	B49049
12	458	82.4	107	2	HP0011
13	458	82.4	130	1	J10079
14	457	82.2	100	2	S29591
15	454	81.7	100	2	S29590
16	454	81.7	107	2	PRO406
17	452	81.3	97	2	S26341
18	446	80.2	106	2	B54378
19	445	80.0	235	2	S35058
20	444	79.9	107	2	S11118
21	442.5	79.6	104	2	JC6076
22	442	79.5	107	2	S11119
23	440	79.1	130	2	A32513
24	437	78.6	97	2	PH1004
25	434	78.1	107	2	S11121
26	433	77.9	107	2	S11117
27	433	77.9	130	2	B32456
28	433	77.9	130	2	S04573
29	430	77.3	99	2	S29585

30	429	77.2	106	2	G27887	Ig kappa chain V r
31	427	76.8	97	2	PH1004	Ig light chain V r
32	426	76.6	94	2	S26440	Ig kappa chain V r
33	426	76.6	99	2	D38601	Ig kappa chain V r
34	425	76.4	109	2	PRO405	Ig light chain V r
35	423	76.1	95	2	D33730	Ig kappa chain V r
36	423	76.1	107	2	S11112	Ig kappa chain V r
37	423	76.1	107	2	S11113	Ig kappa chain V r
38	423	76.1	108	2	S29581	Ig kappa chain V r
39	421	75.7	108	2	PL0278	Ig kappa chain V r
40	420	75.5	109	2	PRO404	Ig light chain pre
41	420	75.5	132	2	S05268	Ig kappa chain V r
42	419	75.4	106	2	PS0070	Ig kappa chain V r
43	418	75.2	107	1	KVMSX4	Ig kappa chain V r
44	417	75.0	108	2	PL0276	Ig kappa chain V r
45	417	75.0	108	2	PL0277	Ig kappa chain V r

## ALIGNMENTS

## RESULT 1

PL0082

Ig kappa chain V region (2D3) - mouse

C:Species: Mus musculus (house mouse)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 21-Jan-2000

C:Accession: PL0082

R:Meek, K.; Hasemann, C.; Pollok, B.; Alkan, S.S.; Brail, M.; Slanov, M.; Urbain, J. Exp. Med. 169, 519-533, 1989

A:Title: Structural characterization of antiidiotypic antibodies: evidence that Ab

A:Reference number: PL0080; MUID:89094246; PMID:2452056

A:Accession: PL0082

A:Molecule type: mRNA

A:Residues: 1-106 <ME>

A:Experimental source: strain BALB/c

A:Note: The sequence shown here is from the V kappa region of an antiidiotypic mon

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotrimer; Immunoglobulin

F:16-89/Domain: Immunoglobulin homology <IMW>

Query Match 94.1% Score 523; DB 2; Length 106;  
Best Local Similarity 94.3% Pred. No. 9.2e-38;  
Matches 100; Conservative 3; Mismatches 3; Indels 0; Gaps

QY 1 QIVLSQSPAILSPGKVTTCRASSVNMHYOQKPPSSPRKATSAISNLASGVPR 60  
DB 1 QIVLSQSPAILSPGKVTTCRASSVNMHYOQKPPSSPRKATSAISNLASGVPR 60

OY 61 FSGSGSGTSTSLTSRVEAEADATYYCOOMSSNPPTFGCGTMLR 106  
DB 61 FSGSGSGTSTSLTSRVEAEADATYYCOOMSSNPPTFGCGTMLR 106

## RESULT 2

A42848

Ig light chain V region - mouse (fragment)

N:Alternate names: L6 anti-tumor antibody

C:Species: Mus musculus (house mouse)

C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000

C:Accession: A42848; S33902

R:Feil, H.P.; Gayle, M.A.; Yelton, D.; Lipsich, L.; Schlawen, G.L.; Marken, J. Biol. Chem. 267, 15552-15558, 1992

A:Title: Chimeric L6 anti-tumor antibody. Genomic construction, expression, and c

A:Reference number: A42848; MUID:92348410; PMID:1635794

A:Accession: A42848

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-107 <FE>

A:Cross-references: EMBL:M90690

A:Note: sequence extracted from NCBI backbone (NCBI:U09558, NCBI:P.109959)

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: Immunoglobulin

F:16-89/Domain: Immunoglobulin homology <IMW>

**THIS PAGE BLANK (USPTO)**

```
DR N-PSDB: AAT79900.  
XX  
PT Inhibiting thrombosis with self-limiting antibody to coagulation  
factor - avoids uncontrolled bleeding by providing only partial  
inhibition  
XX  
PP Example 7; Page 128; 150pp; English.  
XX  
CC This polypeptide comprises a mouse-human chimeric antibody  
light chain in which the variable region is derived from mouse  
anti-human factor IX monoclonal antibody BC2 (see AA624531) and  
human sequences from the immunoglobulin RF-TS3'CL framework.  
CC It can be expressed in transfected mammalian cells utilizing a  
cDNA construct (see AAT79900) obtained by PCR amplification (see  
AAT79997-98) of BC2 cDNA and insertion of the PCR product into  
F9tZHC 1-3 cDNA (see AAT7374). Claimed anti-factor IX chimeric  
antibodies are useful in the treatment of thrombosis.
```

	Sequence	106 AA:
Query Match	94.6%; Score 526;	DB 18; Length 106;
Best Local Similarity	95.3%;	Pred. No. 1,8e+30;
Matches 101;	Conservative 1;	Mismatches 4; Indels 0; Gaps 0.
OY	1 QVLSOSPAILIASAGEKYTMCRASSSYNYMHMYQOKPSPKPMISXSNLASGVPAR 60	
Dd	1 QVLVSOSPAILASAPGEKYTMCRASSSVNVMHMYQOKPGSSPKRPIRYATSNLASGVPAR 60	
OY	61 FSGSGGGTSYSLTISRFEADAAAYTCQQWSSNPFTFGGTMLEIR 106	
Dd	61 FGSAGGSTSYSLTIIRVEADAATYYCQMWSINPRFFGGGTLEIK 106	

```
RESULT 2  
AAU81002  
ID AU81002 standard; Protein; 106 AA.  
XX  
AC AMU81002;  
DT  
XX 09-APR-2002 ((first entry))  
DE Mouse-human light chain polypeptide.  
XX  
KW Human; mouse; BG2; animal post-thromboembolic induced ischaemia;  
KM thrombolytic agent; anti-factor IX antibody; plasminogen activator;  
RM thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;  
KV vasotrophic; cardiac; anti-respiratory syncytial virus;  
XX heavy chain variable region; light chain variable region.  
XX OS Chimeric - Mus sp.  
OS Chimeric - Homo sapiens.  
XX MO2001B7J39-AI.  
XX PD 22-NOV-2001.  
PF 05-OCT-2000; 200OOW-US27438.  
XX PR 15-MAY-2000; 200OUS-0571434.  
XX PA (SMIK ) SMITHLINE BEECHAM CORP.  
PI Barone FC, Blackburn MN, Feuerstein CZ, Toomey JR;  
XX WPI: 2002-082944/11.  
DR N-PSDB: ABK24005.  
XX Treating post-thromboembolic induced ischaemia in an animal by  
administering anti-factor IX antibody in combination with a plasminogen  
activator  
XX Example 7; Page 155-156; 163pp; English.  
XX
```

CC	The invention relates to a method for treating an animal
CC	post-thrombotic induced ischemia or reducing a required dose of a
CC	thrombolytic agent in treatment of an animal post thrombotic induced
CC	ischemia, comprising administering an anti-factor IX antibody or its
CC	fragment, optionally in combination with a plasminogen activator or
CC	thrombolytic agent. The method is useful for treating
CC	post-thrombotic-induced ischemia, for preventing thrombotic
CC	stroke in an animal, and for reducing a required dose of a thrombotic
CC	agent. Sequences AAU80972-AAU81004 represent antibodies and vector
CC	polypeptides used in the method of the invention.
SO	Sequence 106 AA:
DQ	Query Match 94.6%; Score 526; DB 33; Length 106;
DQ	Best Local Similarity 95.3%; Pred. No. 1.8e-33;
DQ	Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0.
OY	1 QIVISOSPALLTSPEKVTMTGRASSSVNWMWVQOKPGSSSEKLSATSNLASGPVAR 60
DB	1 QIVISOSPALLTSPEKVTMTGRASSSVNWMWVQOKPGSSSEKLSATSNLASGPVAR 60
OY	
DB	
OY	61 FSGSGGCTSYLTISRVEADAATYYCOOWSSNPPTFGCGTNIILK 104
DB	61 FSGSGGCTSYLTISRVEADAATYYCOOWSSNPPTFGCGTNIILK 105
OY	
DB	
RESULT 3	
ID	AAW24520 standard; Protein; 107 AA.
AC	AAW24520:
XX	
XX	26-DEC-1997 (first entry)
DE	
XX	Mouse anti-human factor IX antibody BC2 light chain variable region.
KM	Thrombosis: therapy: Factor IX: anticoagulant: monoclonal antibody:
KM	humanised antibody: antibody engineering: light chain: CDR:
KM	complementarity determining region: myocardial infarction:
KM	pulmonary embolism: deep vein thrombosis: coronary angioplasty:
KM	disseminated intravascular coagulation: artificial organ: sepsis:
KM	shunt: prosthesis.
OS	Mus musculus.
FH	Key Location/Qualifiers
FT	Region 24..33
FT	/label=CDR1
FT	/note="(Claim 23)"
FT	Region 49..55
FT	/label=CDR2
FT	/note="(Claim 23)"
FT	Region 88..96
FT	/label=CDR3
FT	/note="(Claim 23)"
PN	MO9726010-A1.
PD	
XX	24-JUL-1997.
XX	
PF	17-JAN-1997; 97WO-US00759.
XX	
PR	24-OCT-1996; 96US-0029119.
PR	17-JAN-1996; 96US-0010108.
PA	(SMIK) SMITHKLINE BEECHAM CORP.
PA	(UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
PI	Blackburn MN, Church WR, Feuerstein GZ, Gross MS;
PI	Nichols AJ, Padian EA, Patel AH, Sylvester DR;
WI	1997-385117/35.

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 27, 2002, 07:25:20 ; Search time 2.25869 Seconds  
(without alignments)  
638.431 Million cell updates/sec

Title: US-09-893-615-1

Perfect score: 15

Sequence: 1 WRMYFSHRHAHLRSP 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR-73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	40.0	296	A40996	phenylalanine 4-mo
2	6	40.0	352	A71447	hypothetical prote
3	6	40.0	556	T19824	hypothetical prote
4	6	40.0	910	E89918	2-oxoglutarate deh
5	6	40.0	942	C81139	probable oxoglutar
6	6	40.0	943	S07776	oxoglutarate dehyd
7	6	40.0	943	G83448	2-oxoglutarate deh
8	6	40.0	1037	T27345	hypothetical prote
9	5	33.3	98	S39402	hupf protein - Bra
10	5	33.3	101	F72515	hypothetical prote
11	5	33.3	142	T22156	hypothetical prote
12	5	33.3	147	D83261	probable transcript
13	5	33.3	148	AB0200	conserved hypotet
14	5	33.3	151	AC2084	conserved hypotet
15	5	33.3	176	T08254	phosphonate metabo
16	5	33.3	190	D70828	probable transposo
17	5	33.3	195	S34993	hypothetical prote
18	5	33.3	196	E81553	nitrite reductase
19	5	33.3	198	S25656	conserved hypotet
20	5	33.3	200	G83032	T-cell surface gly
21	5	33.3	210	T28824	hypothetical prote
22	5	33.3	214	AB3569	hypothetical prote
23	5	33.3	216	A99333	protocatechuate 3,
24	5	33.3	216	D70554	hypothetical prote
25	5	33.3	216	T38520	hypothetical prote
26	5	33.3	217	G69547	conserved hypotet
27	5	33.3	225	T15228	probable histone B
28	5	33.3	226	A11168	transcription regu
29	5	33.3	234	A55367	phosphatase codb -

30	5	33.3	234	2	T34600	hypothetical prote
31	5	33.3	235	1	RWHU78	T-cell surface gly
32	5	33.3	243	2	C95266	probable ABC trans
33	5	33.3	244	2	T29079	hypothetical prote
34	5	33.3	246	2	T31956	hypothetical prote
35	5	33.3	248	2	H95334	probable transcript
36	5	33.3	249	2	T37609	hypothetical zinc-
37	5	33.3	252	2	A84866	probable glyoxalas
38	5	33.3	252	2	AC1181	hypothetical prote
39	5	33.3	252	2	AD1538	hypothetical prote
40	5	33.3	253	2	D75301	hypothetical prote
41	5	33.3	254	2	H83334	probable transcript
42	5	33.3	257	2	T48058	RING-H2 zinc flnge
43	5	33.3	261	2	S51935	probable MADS-box
44	5	33.3	270	2	B84813	probable RING zinc
45	5	33.3	273	2	G83968	hypothetical prote

## ALIGNMENTS

RESULT 1  
A40996 phenylalanine 4-monooxygenase (EC 1.14.16.1) - Chromobacterium violaceum  
N:Alternate names: phenylalaninase, phenylalanine 4-hydroxylase  
C:Species: Chromobacterium violaceum  
C>Date: 03-Aug-1992 #sequence, revision 03-Aug-1992 #text-change 31-Mar-2000  
C:Accession: A40996; B40996  
R:Onishi, A.; Liotta, L.J.; Benkovic, S.J.  
J. Biol. Chem. 266, 18454-18459, 1991  
A:Title: Cloning and expression of Chromobacterium violaceum phenylalanine hydroxylase  
A:Reference number: A40996; MUID:92011593; PMID:1655752  
A:Accession: A40996  
A:Molecule type: DNA  
A:Residues: 1-296 <ONS>  
A:Cross-references: GB:M55915; NID:g144481; PIDN:AAA23115.1; PID:g144482  
A:Accession: B40996  
A:Molecule type: protein  
A:Residues: 1-20 <ONS>  
A:Experimental source: strain ATCC 12540  
C:Keywords: bioplerin; oxidoreductase

Query Match  
Query Local Similarity 40.0%; Score 6; DB 2; Length 296;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RHHLR 13  
Db 192 RHHLR 197

RESULT 2  
A71447 hypothetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: columbia  
C>Date: 03-Aug-1998 #sequence, revision 03-Aug-1998 #text-change 05-Dec-1998  
C:Accession: A71447  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D  
P.; Weiler, H.; Wedler, E.; Wamboldt, R.; Weltzenegger, T.; Pohl, T.M.; Terry, N.; G  
avanagh, T.; Hempel, S.; Kotter, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen  
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.;  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t  
A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: A71447  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-352 <BEV>  
A:Cross-references: GB:297344; NID:g2245126; PID:e327078; PID:g2245129  
C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 40.0%; Score 6; DB 2; Length 352;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AMLRSP 15  
DB 179 AMLRSP 184

RESULT 3

hypothetical protein C38D4.4 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T19824

R:Colles, L.  
submitted to the EMBL Data Library, October 1994

A:Reference number: Z19183

A:Accession: T19824

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-556 <M12>

A:Cross-references: EMBL:Z46241; PIDN:CAA86317.1; GSPDB:GN00021; CESP:C38D4.4

A:Experimental source: clone C38D4

C:Genetics:

A:Gene: CESP:C38D4.4

A:Map position: 3  
A:introns: 40/2; 80/2; 148/3; 257/3; 370/1; 405/1; 488/3

Query Match 40.0%; Score 6; DB 2; Length 556;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AMLRSP 15  
DB 126 AMLRSP 131

RESULT 4

E89918 2-oxoglutarate dehydrogenase E1 [Imported] - *Staphylococcus aureus* (strain N315)

C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: E89918

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc, M.; A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1235-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: E89918

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-910 <KUR>

A:Cross-references: GB:BA000018; PID:q13701210; PIDN:BA842505.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: odha

C:Superfamily: oxoglutarate dehydrogenase (lipoamide); thiamin pyrophosphate-binding dom

Query Match 40.0%; Score 6; DB 2; Length 910;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRHA 10  
DB 599 FSHRHA 604

RESULT 5

C81139

probable oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2) E1 component NMA1149 [lm

C:Species: *Neisseria meningitidis*

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001

C:Accession: C81139; H81881

R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

rt, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignan, V.; Pizze, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: C81139

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-942 <NET>

A:Cross-references: GB:AE002446; GB:AE002098; NID:g7226185; PIDN:AAFA1361.1; PID:g722

A:Experimental source: serogroup B, strain MC58

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: H81881

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-942 <PAR>

A:Cross-references: GB:AL62755; GB:AL157959; NID:g7379742; PIDN:CAB84411.1; PID:g737

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: suca; NMB0955; NMA1149

C:Superfamily: oxoglutarate dehydrogenase (lipoamide); thiamin pyrophosphate-binding

C:Keywords: oxidoreductase

Query Match 40.0%; Score 6; DB 2; Length 942;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRHA 10  
DB 626 FSHRHA 631

RESULT 6

S07776 oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2) - *Azotobacter vinelandii*

N:Alternate names: 2-oxoglutarate dehydrogenase complex chain E1

C:Species: *Azotobacter vinelandii*

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 15-Oct-1999

C:Accession: S07776

R:Schlitz, E.; Westphal, A.H.; Hanemaaijer, R.; de Kok, A.

Eur. J. Biochem. 187, 229-234, 1990

A:Title: The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 1. Mol

A:Reference number: S07776; MUID:50126823; PMID:2404759

A:Accession: S07776

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-943 <SCH>

A:Cross-references: GB:X52433; NID:g39231; PIDN:CAA36680.1; PID:g39232

C:Genetics:

A:Gene: suca

C:Superfamily: oxoglutarate dehydrogenase (lipoamide); thiamin pyrophosphate-binding

C:Keywords: oxidoreductase; thiamin pyrophosphate; tricarboxylic acid cycle

F:359-405/Domain: thiamin pyrophosphate-binding domain homology <PPB>

Query Match 40.0%; Score 6; DB 1; Length 943;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRHA 10  
DB 631 FSHRHA 636

RESULT 7  
G83448  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: G83448  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Micozuchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
adman, S.; Yun, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:2043737; PMID:10984043  
A:Accession: G83448  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-943 <STO>  
A:Cross-references: GB:AE004586; GB:AE004091; NID:99947536; PIDN:AG04974.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: suCA; PA1585  
C:Superfamily: oxoglutarate dehydrogenase (lipoamide); thiamin pyrophosphate-binding dom  
Query Match 40.0%; Score 6; DB 2; Length 943;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 5 FSHRA 10  
Db 631 FSHRA 636

RESULT 8  
T27345  
hypothetical protein Y70C5A.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T27345  
R:Steward, C.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: Z20349  
A:Accession: T27345  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1037 <WIL>  
A:Cross-references: EMBL:Z99282; PIDN:CA16532.1; CESP:Y70C5A.2  
A:Experimental source: clone Y70C5A  
C:Genetics:  
A:Gene: CESP:Y70C5A.2  
A:Introns: 21/2; 71/1; 144/3; 246/1; 283/1; 319/2; 559/3; 636/3; 671/2; 950/3  
Query Match 40.0%; Score 6; DB 2; Length 1037;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 10 AHLRSP 15  
Db 983 AHLRSP 988

RESULT 9  
S39402  
hupf protein - Bradyrhizobium japonicum  
C:Species: Bradyrhizobium japonicum  
C:Date: 07-Oct-1994 #sequence\_revision 26-May-1995 #text\_change 24-Sep-1999  
C:Accession: S39402  
R:van Soom, C.; Browne, J.; Verreth, C.; Vanderleyden, J.  
J. Mol. Biol. 234, 508-512, 1993  
A:Title: Nucleotide sequence analysis of four genes, hupC, hupD, hupE and hupG, downstre  
A:Reference number: S39400; MUID:94047099; PMID:8230232  
A:Accession: S39402  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-98 <VAN>

A:Cross-references: EMBL:Z21948; NID:9311536; PIDN:CA79945.1; PID:9311539  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993  
C:Superfamily: hydrogenase expression/formaton protein hupC  
Query Match 33.3%; Score 5; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 10 AHLRS 14  
Db 93 AHLRS 97

RESULT 10  
F72515  
hypothetical protein APE2099 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: F72515  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: F72515  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-101 <KAW>  
A:Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BA81110.1; PID:95105798  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE2099  
C:Superfamily: Aeropyrum pernix hypothetical protein APE2099  
Query Match 33.3%; Score 5; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 10 AHLRS 14  
Db 2 AHLRS 6

RESULT 11  
T22156  
hypothetical protein F44A6.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T22156  
R:Stinson, J.  
submitted to the EMBL Data Library, August 1995  
A:Reference number: Z19524  
A:Accession: T22156  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-142 <WIL>  
A:Cross-references: EMBL:Z50858; PIDN:CAA90724.1; GSPDB:GN00028; CESP:F44A6.4  
A:Experimental source: clone F44A6  
C:Genetics:  
A:Gene: CESP:F44A6.4  
A:Map position: X  
A:Introns: 17/1; 41/1; 97/1  
C:Superfamily: Caenorhabditis elegans hypothetical protein F44A6.4  
Query Match 33.3%; Score 5; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 7 HRHAH 11  
Db 4 HRHAH 8

RESULT 12  
D83261  
Probable transcription regulator PA3067 [Imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83261  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lardig, K.; Lam,  
N.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A:Reference number: AB2950; M0ID:20437337; PMID:10984043  
A:Accession: D83261  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-147 <STO>  
A:Cross-references: GB:AE004731; GB:AE004091; NID:g9949171; PIDN:AAG06455.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Superfamily: Mycobacterium leprae hypothetical protein MLCB57.31

Query Match 33.3%; Score 5; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRS 14  
DB 11 AHLRS 15

RESULT 13  
AB0200  
Conserved hypothetical protein YPO1639 [Imported] - *Yersinia pestis* (strain CO92)  
C:Species: *Yersinia pestis*  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
C:Accession: AB0200  
R:parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.  
demparrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A:Reference number: AB0001; M0ID:21470413; PMID:11586360  
A:Accession: AB0200  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-148 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC90461.1; PID:g15979677; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO1639  
C:Superfamily: unassigned mutr domain proteins; mutr domain homology

Query Match 33.3%; Score 5; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 HLRSR 15  
DB 117 HLRSR 121

RESULT 14  
AC2084  
phosphonate metabolism protein [Imported] - *Nostoc* sp. (strain PCC 7120)  
C:Species: *Nostoc* sp.  
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AC2084  
R:kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*  
A:Reference number: AB1807; M0ID:21595285; PMID:11759840

A:Accession: AC2084  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-151 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA073924.1; PID:g17131316; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: phnG  
C:Superfamily: *Escherichia coli* phnG protein

Query Match 33.3%; Score 5; DB 2; Length 151;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 SHRHA 10  
DB 91 SHRHA 95

RESULT 15  
T08254  
Probable transposase H0455 - *Halobacterium* sp. (strain NRC-1) insertion sequence ISH5  
N:Alternate names: hypothetical protein H1903  
C:Species: *Halobacterium* sp.  
A:Variety: strain NRC-1  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 31-Jan-2000  
C:Accession: T08254; T08389  
R:Ng, W.V.; Clufo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall, B  
Genome Res. 8, 1131-1141, 1998  
A:Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid o  
A:Reference number: Z16408; M0ID:99063795; PMID:9847077  
A:Accession: T08254  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-176 <NGM>  
A:Cross-references: EMBL:AF016485; NID:g2822278; PID:g2822315; HALOSP:H0455  
A:Experimental source: strain NRC-1  
A:Genetics: COP1  
A:Accession: T08389  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-176 <DAS>  
A:Cross-references: EMBL:AF016485; NID:g2822278; PID:g2822450; HALOSP:H1903  
A:Experimental source: strain NRC-1  
A:Genetics: COP2  
C:Genetics: <COP1>  
A:Gene: HALOSP:H0455  
A:Genome: plasmid pNRC100  
C:Genetics: <COP2>  
A:Gene: HALOSP:H1903  
A:Genome: plasmid pNRC100  
A:Mobile element: insertion sequence ISH5

Query Match 33.3%; Score 5; DB 2; Length 176;  
Best Local Similarity 100.0%; Pred. No. 1e+07;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RHABL 12  
DB 124 RHABL 128

Search completed: November 27, 2002, 07:31:16  
Job time : 8.25869 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:23:05 ; Search time 1.44788 Seconds  
(without alignments) 429.695 Million cell updates/sec

Title: US-09-893-615-1  
Perfect score: 15  
Sequence: 1 WRMYFSHRHRLRSP 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues  
Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SWISSPROT\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	6	40.0	943	1	ODOL_AZOVI
2	5	33.3	98	1	HUPF_BRAJA
3	5	33.3	162	1	YUJG_RHISN
4	5	33.3	189	1	UREF_STAXY
5	5	33.3	198	1	CD8A_PONPY
6	5	33.3	205	1	FRDA_MOUSE
7	5	33.3	235	1	CD8A_HUMAN
8	5	33.3	309	1	VAGE_ECOLI
9	5	33.3	312	1	V531_METTH
10	5	33.3	331	1	GL2M_ARATH
11	5	33.3	349	1	AMBP_RAT
12	5	33.3	372	1	Y906_MYCTU
13	5	33.3	377	1	RAPB_BACSU
14	5	33.3	384	1	Y823_MYCLE
15	5	33.3	388	1	GSPE_ABRHY
16	5	33.3	394	1	UXUA_HAEIN
17	5	33.3	395	1	STTB_MOUSE
18	5	33.3	397	1	O22A_DROME
19	5	33.3	397	1	O22B_DROME
20	5	33.3	410	1	APKA_ARATH
21	5	33.3	411	1	VGLG_HSVBR
22	5	33.3	411	1	VGLX_HSVBR
23	5	33.3	416	1	PYRC_DEIRA
24	5	33.3	426	1	YEY2_YEAST
25	5	33.3	443	1	GLMU_BUCAT
26	5	33.3	459	1	SYN_BUCAT
27	5	33.3	466	1	TRP2_ARATH
28	5	33.3	475	1	KG3A_HUMAN
29	5	33.3	483	1	KG3A_RAT
30	5	33.3	490	1	DNMA_CAUCR
31	5	33.3	494	1	VGA_BPAL3
32	5	33.3	494	1	VGA_BPAL3
33	5	33.3	494	1	VGA_BPAL3

34	5	33.3	505	1	FIXL_RHIME	P10955 rhizobium m
35	5	33.3	508	1	GALT_BACAD	O9KdV2 bacillus ha
36	5	33.3	523	1	SYN_TREPA	O83618 treponema p
37	5	33.3	545	1	CIP4_HUMAN	O15642 homo sapien
38	5	33.3	571	1	PUR6_YEAST	P12164 saccharomyc
39	5	33.3	628	1	V7OK_TYMW	P10357 turnip yell
40	5	33.3	628	1	V7OK_TYMW	P20131 turnip yell
41	5	33.3	642	1	NOGL_SCHRO	O94659 schizosacch
42	5	33.3	647	1	NOGL_YEAST	O02892 saccharomyc
43	5	33.3	699	1	YQOA_CAEEL	O09299 caenorhabdi
44	5	33.3	701	1	TBX2_MOUSE	O60707 mus musculu
45	5	33.3	702	1	TBX2_HUMAN	O13207 homo sapien

## ALIGNMENTS

```

RESULT 1
ID      ODOL_AZOVI      STANDARD:      PRT:      943 AA.
AC      P20707
DT      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2) (Alpha-
DE      ketoglutarate dehydrogenase).
GN      SUCR OR ODBA.
OS      Azotobacter vinelandii.
OC      Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC      Azotobacter.
OX      NCBI_TaxID=354;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90126823; PubMed=2404759;
RA      Schulze E., Westphal A.H., Hanemaaijer R., de Kok A.;
RT      "The 2-oxoglutarate dehydrogenase complex from Azotobacter
RT      vinelandii. 1. Molecular cloning and sequence analysis of the gene
RT      encoding the 2-oxoglutarate dehydrogenase component."
RL      Eur. J. Biochem. 187:229-234(1990).
RN      [2]
RP      SEQUENCE OF 909-943 FROM N.A.
RX      MEDLINE=90126825; PubMed=2404760;
RA      Westphal A.H., de Kok A.;
RT      "The 2-oxoglutarate dehydrogenase complex from Azotobacter
RT      vinelandii. 2. Molecular cloning and sequence analysis of the gene
RT      encoding the succinyltransferase component."
RL      Eur. J. Biochem. 187:235-239(1990).
CC      -!- FUNCTION: THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX CATALYZES THE
CC      OVERALL CONVERSION OF 2-OXOGLUTARATE TO SUCCINYL-COA & CO(2). IT
CC      CONTAINS MULTIPLE COPIES OF 3 ENZYMIC COMPONENTS: 2-OXOGLUTARATE
CC      DEHYDROGENASE (E1), DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE (E2) AND
CC      LIPOAMIDE DEHYDROGENASE (E3).
CC      -!- CATALYTIC ACTIVITY: 2-oxoglutarate + lipoamide = S-
CC      succinylidihydrolipoamide + CO(2).
CC      -!- COFACTOR: THIAMINE PYROPHOSPHATE.
CC      -!- SUBUNIT: HOMODIMER.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X52433; CA36680.1; -.
DR      EMBL: X52432; CA36677.1; -.
DR      PIR: S07776; S07776.
DR      InterPro: IPR001017; Dehydrogenase_E1.
DR      InterPro: IPR000360; Transketolase.
DR      Pfam: PF00676; E1-dehydrog. 1.
DR      Pfam: PF02779; transket.pyr. 1.
DR      TIGRfams: TIGR00239; 2oxo_dh_E1; 1.

```

KM Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate.  
SQ SEQUENCE 943 AA; 105687 MW; D3F35356D454E2A1 CRC64;

Query Match 40.0%; Score 6; DB 1; Length 943;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRHA 10  
|||||  
Db 631 FSHRHA 636

## RESULT 2

HUPE\_BRAJA STANDARD; PRT; 98 AA.

AC 045252;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE Hydrogenase expression/formation protein hupf.  
GN HUPE

OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OX NCBI\_TaxID=375;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=94047099; PubMed=8230232;  
RA van Soom C., Browneys J., Verreth C., Vanderleyden J.,  
RT "Nucleotide sequence analysis of four genes, hupC, hupD, hupF and  
hupG, downstream of the hydrogenase structural genes in  
Bradyrhizobium japonicum."

RT Bradyrhizobium japonicum.  
RL J. Mol. Biol. 234:508-512(1993).  
CC -1- SIMILARITY: BELONGS TO THE HUPE/HYPC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: 221948; CAA79945.1; -  
DR InterPro: IPR001109; HupF\_HYPC.  
DR Pfam: PF01455; HupF\_HYPC.1.  
DR ProDom: PD003312; HupF\_HYPC.1.  
DR TIGRFAMs: TIGR00074; hupC\_hupF.1.  
DR PROSITE: PS01097; HupF\_HYPC; FALSE\_NEG.  
SQ SEQUENCE 98 AA; 10633 MW; A9BD353F9D4240F7 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 98;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRS 14  
|||||  
Db 93 AHLRS 97

## RESULT 3

Y4JG\_RHISN STANDARD; PRT; 162 AA.

AC P55507;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE Hypothetical 17.9 kDa protein Y4JG.  
GN Y4JG.

OS Rhizobium sp. (strain NGR234).  
OC Plasmid sym pNCR234a.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.

OX NCBI\_TaxID=394;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97305956; PubMed=9163424;  
RA Freidberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,  
RA Perret X.,  
RT "Molecular basis of symbiosis between Rhizobium and legumes."

RL Nature 387:394-401(1997).  
CC -1- SIMILARITY: NONE OBVIOUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE000079; AAB91719.1; -  
KM Hypothetical protein; Plasmid; Transmembrane.  
FT TRANSMEM 108 126

SQ SEQUENCE 162 AA; 17910 MW; B32FA1D10F64950C CRC64;

Query Match 33.3%; Score 5; DB 1; Length 162;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRS 14  
|||||  
Db 22 AHLRS 26

## RESULT 4

UREF\_STAXY STANDARD; PRT; 189 AA.

AC P42876;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Urease accessory protein uref.

GN UREF.  
OS Staphylococcus xylosum.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_TaxID=1288;

RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 20267 / Isolate C2A;  
RA Jose J.,  
RL Thesis (1994), University of Saarlandes, Germany.

CC -1- FUNCTION: PROBABLY FACILITATES NICKEL INCORPORATION.  
CC -1- SIMILARITY: BELONGS TO THE UREF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z35136; CAA84508.1; -

DR InterPro: IPR002639; Uref.

DR Pfam: PF01730; Uref.1.  
DR ProDom: PD004961; Uref.1.

KW NICKEL.  
SQ SEQUENCE 189 AA; 21987 MW; 6E6965557014FA6F CRC64;

Query Match 33.3%; Score 5; DB 1; Length 189;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 AHLRS 13  
|||||

Db 4 HAHLE 8

RESULT 5  
CD8A\_PONPY STANDARD; PRT; 198 AA.

AC P30433;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE T-cell surface glycoprotein CD8 alpha chain precursor (T-lymphocyte differentiation antigen T8/LEU-2).  
GN CD8A.  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pongo.  
OX NCBI\_TaxID=9600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Isolate Jarl;  
RX MEDLINE=92307742; PubMed=1612644;  
RA Lawlor D.A., Parham P.;  
RT "Structure of CD8 alpha and beta chains of the orangutan: novel patterns of mRNA splicing encoding hingeless polypeptides.";  
RL Immunogenetics 36:121-125(1992).  
CC -1- FUNCTION: IDENTIFIES CYTOTOXIC/SUPPRESSOR T-CELLS THAT INTERACT WITH MHC CLASS I BEARING TARGETS. CD8 IS THOUGHT TO PLAY A ROLE IN THE PROCESS OF T-CELL MEDIATED KILLING. CD8 ALPHA CHAINS BINDS TO CLASS I MHC MOLECULES ALPHA-3 DOMAINS.  
CC -1- SUBUNIT: IN GENERAL HETERODIMER OF AN ALPHA AND A BETA CHAIN LINKED BY TWO DISULFIDE BONDS. CAN ALSO FORM HOMODIMERS.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X60223; CAA42784.1; -.  
DR HSSP: P01732; LCD8.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig\_1.  
DR SMART: SM00406; IgV\_1.  
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell; Immune response; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 198  
FT -----  
FT DOMAIN 22 145  
FT TRANSMEM 146 166  
FT DOMAIN 167 198  
FT DOMAIN 22 135  
FT DISULFID 43 115  
SQ SEQUENCE 198 AA; 22099 MW; F3EC093EAD805561 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 198;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MYFSH 7  
Db 123 MYFSH 127

RESULT 6  
FRDA\_MOUSE STANDARD; PRT; 207 AA.  
ID FRDA\_MOUSE

AC O35943;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Frataxin, mitochondrial precursor (Fxn).  
GN FRDA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.  
RX MEDLINE=97385237; PubMed=9241270;  
RA Koutnikova H., Campuzano V., Fourny F., Dolle P., Cazallani O., Koenig M.;  
RT "Studies of human, mouse and yeast homologues indicate a mitochondrial function for frataxin.";  
RL Nat. Genet. 16:345-351(1997).  
CC -1- FUNCTION: PROBABLY INVOLVED IN IRON HOMEOSTASIS.  
CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
CC -1- PTM: PROCESSED BY MITOCHONDRIAL PROCESSING PEPTIDASE (MPP).  
CC -1- SIMILARITY: BELONGS TO THE FRATAXIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: U95736; AAB67778.1; -.  
DR HSSP: Q16595; IDLX.  
DR MGD: MGI:1096879; Frda.  
DR InterPro: IPR001794; Frataxin.  
DR InterPro: IPR002908; Frataxin\_1like.  
DR Pfam: PF01491; Frataxin\_Cyay; 1.  
DR PRINTS: PR00904; FRATAXIN.  
DR PRODOM: PD006646; Frataxin\_1like; 1.  
DR PROSITE: PS01344; FRATAXIN\_1; 1.  
DR PROSITE: PS50810; FRATAXIN\_2; 1.  
KW Mitochondrion; Transit peptide.  
FT TRANSIT 1 40  
FT CHAIN 41 207  
SQ SEQUENCE 207 AA; 22924 MW; C46FD1B44FB26A2 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 207;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RHAHL 12  
Db 51 RHAHL 55

RESULT 7  
CD8A\_HUMAN STANDARD; PRT; 235 AA.  
ID CD8A\_HUMAN  
AC P01732;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE T-cell surface glycoprotein CD8 alpha chain precursor (T-lymphocyte differentiation antigen T8/Leu-2).  
GN CD8A OR MAL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85099337; PubMed=3871356;  
RA Litman D.R., Thomas Y., Madden P.J., Chess L., Axel R.;

RT "The isolation and sequence of the gene encoding T8: a molecule  
RT defining functional classes of T lymphocytes.";  
RL Cell 40:237-246(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86103103; PubMed=3936473;  
RA Parnes J.R., Sizer K.C., Sukhatme V.P., Hunkapiller T.;  
RT "Structure of Leu-2/8 as deduced from the sequence of a cDNA clone";  
RL Behring Inst. Mitt. 77:48-55(1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85124610; PubMed=3918796;  
RA Sukhatme V.P., Sizer K.C., Volimer A.C., Hunkapiller T.,  
RA Parnes J.R.;  
RT "The T cell differentiation antigen Leu-2/8 is homologous to  
RT immunoglobulin and T cell receptor variable regions.";  
RL Cell 40:591-597(1985).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90035142; PubMed=2509342;  
RA Nakayama K.-I., Tokito S., Okumura K., Nakauchi H.;  
RT "Structure and expression of the gene encoding CD8 alpha chain (Leu-  
RT 2/8).";  
RL Immunogenetics 30:393-397(1989).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89215302; PubMed=2496167;  
RA Norment A.M., Lonberg N., Lacey E., Littman D.R.;  
RT "Alternatively spliced mRNA encodes a secreted form of human CD8  
RT alpha. Characterization of the human CD8 alpha gene.";  
RL J. Immunol. 142:3312-3319(1989).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 22-135.  
RX MEDLINE=92191292; PubMed=1547508;  
RA Leahy D.J., Axel R., Hendrickson W.A.;  
RT "Crystal structure of a soluble form of the human T cell coreceptor  
RT CD8 at 2.6-A resolution.";  
RL Cell 68:1145-1162(1992).  
CC -1- FUNCTION: IDENTIFIES CYTOTOXIC/SUPPRESSOR T-CELLS THAT INTERACT  
CC WITH MHC CLASS I BEARING TARGETS. CD8 IS THOUGHT TO PLAY A ROLE IN  
CC THE PROCESS OF T-CELL MEDIATED KILLING. CD8 ALPHA CHAINS BINDS TO  
CC CLASS I MHC MOLECULES ALPHA-3 DOMAINS.  
CC -1- SUBUNIT: IN GENERAL HETERODIMER OF AN ALPHA AND A BETA CHAIN  
CC LINKED BY TWO DISULFIDE BONDS. CAN ALSO FORM HOMODIMERS.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: VARIOUS PATTERNS OF DIFFERENTIAL SPLICING  
CC OF CD8 ALPHA TRANSCRIPTS INVOLVE EXCISION OF THE TRANSMEMBRANE OR  
CC CYTOPLASMIC DOMAINS.  
CC -1- PTM: ALL OF THE FIVE MOST CARBOXYL-TERMINAL CYSTEINES ARE USED TO  
CC FORM INTER-CHAIN DISULFIDE BONDS IN DIMERS AND HIGHER MULTIMERS,  
CC WHILE THE FOUR AMINO-TERMINAL CYSTEINES ARE NOT (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.  
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD8a entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd8alpha.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: M26315; AAA79217.1; -  
DR EMBL: M26313; AAA79217.1; JOINED.  
DR EMBL: M26314; AAA79217.1; JOINED.  
DR EMBL: M12824; AAA61133.1; -  
DR EMBL: M12828; AAB04637.1; -  
DR EMBL: M27161; AAA59674.1; -  
DR PIR: A01999; RWHUT8.  
DR PIR: A23824; A22824.  
DR PIR: JP0105; JP0105.

DR PIR: A30604; A30604.  
DR PIR: A45888; A45888.  
DR PDB: 1CD8; 31-JAN-94.  
DR Genew; HGNC:1706; CD8A.  
DR MIM; 186910; -  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig\_1.  
DR SMART; SM00406; IgV\_1.  
KW Immunoglobulin domain; Transmembrane; Glycoprotein; Phosphorylation;  
KW T-cell; Immune response; Signal; 3D-structure; Alternative splicing.  
FT SIGNAL 1 21  
FT CHAIN 22 235  
FT DOMAIN 22 182  
FT TRANSMEM 183 203  
FT DOMAIN 204 235  
FT DOMAIN 22 135  
FT DISULFID 43 115  
FT STRAND 24 27  
FT TURN 35 36  
FT STRAND 39 45  
FT STRAND 54 59  
FT STRAND 68 73  
FT STRAND 79 80  
FT TURN 82 83  
FT TURN 86 88  
FT STRAND 89 94  
FT TURN 95 96  
FT STRAND 97 102  
FT HELIX 107 109  
FT STRAND 111 119  
FT TURN 120 121  
FT STRAND 122 125  
FT STRAND 129 131  
SQ SEQUENCE 235 AA; 25729 MW; FCCA29BA73726BB CRC64;  
Query Match 33.3%; Score 5; DB 1; Length 235;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 MYFSH 7  
Db 123 MYFSH 127  
YAGE\_ECOLI  
ID YAGE\_ECOLI STANDARD; PRT; 309 AA.  
AC P75682; Q9R2D5;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein yage.  
GN YAGE OR B0268.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blatner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [2]  
RP SEQUENCE OF 1-128 FROM N.A.  
RC STRAIN=K12 / WJ110;  
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,

RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,  
 RA Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mochuchi K.,  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of the  
 RL 4.0 - 6.0 mln (189,987 - 281,416bp) region."  
 CC Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE DHPS FAMILY. STRONG, TO E.COLI YJHH.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE000134; AAC73371.1; -  
 DR EMBL: D83536; BAA77934.1; -  
 DR EcGene: EG13344; YAGE.  
 DR InterPro: IPR002220; DHPS.  
 DR Pfam: PF00701; DHPS.1.  
 DR PRINTS: PR00146; DHPICNTHASE.  
 DR PRODOM: PD001859; DHPS.1.  
 DR PROSITE: PS00665; DHPS.1; 1.  
 DR PROSITE: PS00666; DHPS.2; 1.  
 KM Hypothetical protein; lyase; Complete proteome.  
 FT ACT\_SITE 174 174  
 FT CONFLICT 105 106  
 SQ SEQUENCE 309 AA; 3316 MW; 94FC34AD8C852BE CRC64;  
 Query Match 33.3%; Score 5; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 HLRS 14  
 Db 181 HLRS 185  
 RESULT 9  
 Y31\_METH STANDARD; PRT; 312 AA.  
 ID Y31\_METH  
 AC 026631;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MTH531.  
 GN MTH531.  
 OS Methanobacterium thermoautotrophicum.  
 CC Archaea: Euryarchaeota: Methanobacteria: Methanobacteriales:  
 CC Methanobacteriaceae: Methanobacter.  
 CC NCBI\_Taxid=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Qiu D.,  
 RA Spadator R., Vitcare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Pirbhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics."  
 RL J. Bacteriol. 179:7135-7153(1997).  
 CC -1- SIMILARITY: BELONGS TO THE MURDEF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE000836; AAB85037.1; -  
 DR InterPro: IPR000713; Mur\_ligase.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR Pfam: PF01225; Mur\_ligase.1.  
 KM Hypothetical protein; ligase; ATP-binding; Complete proteome.  
 FT NP\_BIND 112 118  
 SQ SEQUENCE 312 AA; 34742 MW; 4AA7AFEF0A793002 CRC64;  
 Query Match 33.3%; Score 5; DB 1; Length 312;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 11 HLRS 15  
 Db 303 HLRS 307  
 RESULT 10  
 GL2M\_ARATH STANDARD; PRT; 331 AA.  
 ID GL2M\_ARATH  
 AC 024495; 024494; 022857;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hydroxyacylglytathione hydrolase, mitochondrial precursor (EC 3.1.2.6)  
 DE (Glyoxalase II) (GLX II).  
 GN GLX2-1 OR AT2G43430 OR T1024.17.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 CC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
 CC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:  
 CC Eustrods II; Brassicales; Brassicaceae: Arabidopsis.  
 CC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-cv. Massilewskij;  
 RX MEDLINE=98009983; PubMed=9349270;  
 RA Maltl M.K., Krishnasamy S., Owen H.A., Makaroff C.A.;  
 RT "Molecular characterization of glyoxalase II from Arabidopsis  
 RT thaliana."  
 RL Plant Mol. Biol. 35:471-481(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-cv. Columbia.  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,  
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,  
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Greasy T.H.,  
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
 RA Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana."  
 RL Nature 402:761-768(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-cv. Columbia;  
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
 RT Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
 RT SSP consortium (Salk/Stanford/PGSC)."  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THIOLESTERASE THAT CATALYZES THE HYDROLYSIS OF S-D-  
 CC LACTOYL-GLUTATHIONE TO FORM GLUTATHIONE AND D-LACTIC ACID.  
 CC -1- CATALYTIC ACTIVITY: (S)-(2-hydroxyacetyl)glutathione + H(2)O -  
 CC -1- COFACTOR: BINDS TWO ZINC IONS (BY SIMILARITY).  
 CC -1- PATHWAY: Glyoxal pathway.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- SIMILARITY: BELONGS TO THE GLYOXALASE II FAMILY.  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: U90927; AAC49865.1; -  
DR EMBL: U90928; AAC49866.1; -  
DR EMBL: AC002335; AAB64315.2; -  
DR EMBL: AY091278; AAM44217.1; -  
DR EMBL: AY063806; AAL36162.1; -  
DR HSSP: Q16775; IOH5.  
DR InterPro: IPR001279; Bactemase-like.  
DR Pfam: PF00753; Lactamase\_B.1.  
KW Hydrolase; Zinc; Mitochondrion; Transit peptide.  
FT TRANSIT 1 76 MITOCHONDRION (POTENTIAL).  
FT CHAIN 77 331 HYDROXYACETYLGLUTATHIONE HYDROLASE.  
FT METAL 131 131 ZINC 1 (BY SIMILARITY).  
FT METAL 133 133 ZINC 1 (BY SIMILARITY).  
FT METAL 135 135 ZINC 2 (BY SIMILARITY).  
FT METAL 136 136 ZINC 2 (BY SIMILARITY).  
FT METAL 189 189 ZINC 1 (BY SIMILARITY).  
FT METAL 208 208 ZINC 1 AND 2 (BY SIMILARITY).  
FT CONFLICT 159 159 D -> H (IN REF. 1; AAC49866).  
SQ SEQUENCE 331 AA; 36499 MW; 2EDCC21B4902419C5 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 331;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 AHLRS 14  
|||||  
Db 274 AHLRS 278

RESULT 11  
AMB\_P\_RAT STANDARD: PRT: 349 AA.  
ID AMBP\_RAT 064240; O63336; P19603;  
AC 01-FEB-1991 (Rel. 17, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE AMBP protein precursor [Contains: Alpha-1-microglobulin; Inter-alpha-trypsin inhibitor light chain (ITR-LC) (Bikunin) (HI-30); Trypsatin].  
GN AMBP OR ITRL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=92182014; PubMed=1371936;  
RA Lindqvist A., Bratt T., Altieri M., Kasterstroem B.;  
RT "Rat alpha 1-microglobulin: co-expression in liver with the light chain of inter-alpha-trypsin inhibitor.";  
RL Biochim. Biophys. Acta 1130:63-67(1992).  
RN [2]  
RN SEQUENCE OF 141-195 FROM N.A.  
RX MEDLINE=87033744; PubMed=2429963;  
RA Kasterstroem M., Bjorck L., Akerstroem B.;  
RT "RNA in the rat.";  
RL J. Biol. Chem. 261:15070-15074(1986).  
RN [3]  
RN SEQUENCE OF 283-343, AND CHARACTERIZATION.  
RC STRAIN=WiStar;  
RX MEDLINE=89053978; PubMed=3263966;  
RA Kido H., Yokogoshi Y., Katunuma N.;  
RT "Kunitz-type protease inhibitor found in rat mast cells. Purification, properties, and amino acid sequence.";

RL J. Biol. Chem. 263:18104-18107(1988).  
RN [4]  
RP PROCESSING.  
RX MEDLINE=9414892; PubMed=7508921;  
RA Itoh H., Ide H., Ishikawa N., Nawa Y.;  
RT "Mast cell protease inhibitor, trypstatin, is a fragment of inter-alpha-trypsin inhibitor light chain.";  
RL J. Biol. Chem. 269:3818-3822(1994).  
CC -1- FUNCTION: ALPHA-1-MICROGLOBULIN OCCURS IN MANY PHYSIOLOGICAL FLUIDS INCLUDING PLASMA, URINE, AND CEREBROSPINAL FLUID. IT APPEARS NOT ONLY AS A FREE MONOMER BUT ALSO IN COMPLEXES WITH IGA AND ALBUMIN (BY SIMILARITY).  
CC -1- FUNCTION: INTER-ALPHA-TRYPsin INHIBITOR. PRESENT IN PLASMA AND URINE. INHIBITS TRYPSIN, PLASMIN, AND LYSOSOMAL GRANULOCYTIC ELASTASE (BY SIMILARITY).  
CC -1- FUNCTION: Trypsatin is a trypsin inhibitor. It inhibits blood coagulation factor Xa and trypsin about 100-fold more rapidly than porcine pancreatic trypsin and chymase. It is a monomer but is also found in mast cells as a complex with trypase.  
CC -1- SUBUNIT: I-ALPHA-I plasma protease inhibitors are assembled from one or two heavy chains (H1, H2 or H3) and one light chain, bikunin. Inter-alpha-inhibitor (I-ALPHA-I) is composed of H1, H2 and bikunin. Inter-alpha-like inhibitor (I-ALPHA-LI) of H2 and bikunin, and pre-alpha-inhibitor (P-ALPHA-I) of H3 and bikunin (by similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Synthesized in the liver and secreted in plasma. Trypsatin is present in mast cell granules.  
CC -1- PTM: The precursor is proteolytically processed into two separately functioning proteins.  
CC -1- PTM: Alpha-1-microglobulin contains covalently linked brown-yellow chromophores (by similarity).  
CC -1- PTM: Heavy chains are interlinked with bikunin via a chondroitin 4-sulfate bridge to the their C-terminal aspartate (by similarity).  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE LIPOCALIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 2 BPTI/KUNITZ INHIBITOR DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

```

FT BINDING 110 110 CHROMOPHORE (BY SIMILARITY).
FT BINDING 136 136 CHROMOPHORE (BY SIMILARITY).
FT BINDING 148 148 CHROMOPHORE (BY SIMILARITY).
FT DISULFID 90 187 BY SIMILARITY.
FT DISULFID 230 280 BY SIMILARITY.
FT DISULFID 239 263 BY SIMILARITY.
FT DISULFID 255 276 BY SIMILARITY.
FT DISULFID 286 336 BY SIMILARITY.
FT DISULFID 295 319 BY SIMILARITY.
FT DISULFID 311 332 BY SIMILARITY.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 INHIBITORY SITE (P1) (CHYMOTRYPSIN,
FT ACT_SITE 240 241 ELASTASE) (BY SIMILARITY).
FT ACT_SITE 296 297 INHIBITORY SITE (P1) (TRYPSIN) (BY
FT ACT_SITE 142 142 SIMILARITY).
FT CONFLICT 302 302 G -> A (IN REF. 2).
FT CONFLICT 323 323 W -> L (IN REF. 3).
FT CONFLICT 330 331 G -> N (IN REF. 3).
FT CONFLICT 334 334 KE -> PK (IN REF. 3).
SO SEQUENCE 349 AA: 38851 MW: 18787DCB0824E01 CRC64:

Query Match 33.3%; Score 5; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRH 9
Db 137 FSHRH 141

RESULT 12
Y906_MYCTU STANDARD: PRT: 372 AA.
ID Y906_MYCTU
AC Q10562;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV0906 precursor.
GN RV0906 OR MT0929 OR MTCT31.34.
OS Mycobacterium tuberculosis.
OC Bacteria: Actinobacteria: Actinobacteria (class): Actinobacteridae:
OC Actinomycetales: Corynebacteriaceae: Mycobacterium.
OX NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gass S., Barry C.E. III, Tekela F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hovnsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Flatschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Kelson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishal W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO K.PNEUMONIAE ROMA.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z73101; CA97381.1; -
DR EMBL: AE006979; AAK45176.1; -
DR TIGR: MT0929; -
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 372 HYPOTHEICAL PROTEIN RV0906.
SO SEQUENCE 372 AA: 40641 MW: 0A85549D2429D335 CRC64:

Query Match 33.3%; Score 5; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRS 14
Db 195 AHLRS 199

RESULT 13
RAPB_BACSU STANDARD: PRT: 377 AA.
ID RAPB_BACSU
AC P70962;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Response regulator aspartate phosphatase B (EC 3.1.-.-) (Stage 0
DE sporulation protein P).
GN RAPB OR SPOOP.
OS Bacillus subtilis.
OC Bacteria: Firmicutes: Bacillales: Bacillaceae: Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=95094268; PubMed=8001132;
RA Perez M., Hanstein C., Welsh K., Djavahishvili T., Glaser P.,
RA Hoch J.;
RT "Multiple protein-aspartate phosphatases provide a mechanism for the
RT integration of diverse signals in the control of development in B.
RT subtilis."
RL Cell 79:1047-1055(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertaino M.G., Bessieres P., Bolojin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hujilo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Melledo R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilwa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porrolik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,

```

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takenchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambolt R., Wedler E., Wedler H., Wellemsger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Nature 390:249-256(1997).  
 CC -1- FUNCTION: Prevents sporulation by dephosphorylating SpoOF.  
 CC -1- SIMILARITY: BELONGS TO THE RAP FAMILY.  
 CC -1- SIMILARITY: CONTAINS 6 TPR REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 281356; CAB03684.1; -;  
 DR EMBL: 291122; CAB15686.1; -;  
 DR Subtilist; BG11965; rapB.  
 DR InterPro: IPR001440; TPR.  
 DR Pfam: PF00515; TPR: 4.  
 KM Hydrolyase; Repeat; TPR repeat; Sporulation; Complete proteome.  
 FT REPEAT 100 133 TPR 1.  
 FT REPEAT 147 180 TPR 2.  
 FT REPEAT 181 214 TPR 3.  
 FT REPEAT 221 254 TPR 4.  
 FT REPEAT 260 293 TPR 5.  
 FT REPEAT 335 368 TPR 6.  
 SQ SEQUENCE 377 AA; 45037 MW; 617B29614B483971 CRC64;  
 Query Match 33.3%; Score 5; DB 1; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 4 YESHR 8  
 Db 344 YESHR 348  
 RESULT 14  
 Y823\_MYCLE STANDARD; PRT; 384 AA.  
 ID 050049;  
 AC 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein ML2186 (NCRB).  
 GN ML2186.  
 OS Mycobacterium leprae.  
 OC Bacteria: Actinobacteria: Actinobacteria (class): Actinobacteridae;  
 OC Actinomycetales: Corynebacterineae: Mycobacteriaceae: Mycobacterium.  
 OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Smith D.R., Rodison K.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Elgimeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Whangell K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Davis K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holtroyd S., Horsley T., Jagsels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,

RA Rutter S., Seeger K., Simon S., Simmonds M., Skellon J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrett B.G.;  
 RT "Massive gene decay in the leprosy bacillus.";  
 RL Nature 409:1007-1011(2001).  
 CC -1- SIMILARITY: BELONGS TO THE UPF0034 (NIF3/SNM1) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U15182; AAA62994.1; -;  
 DR EMBL: AL583824; CAC31141.1; -;  
 DR Lepioma; ML2186; -;  
 DR InterPro: IPR003009; FMN\_enzyme.  
 DR InterPro: IPR004652; NIF3\_YndG.  
 DR InterPro: IPR001269; UPF0034.  
 DR Pfam: PF01207; UPF0034: 1.  
 DR TIGRfam: TIGR00737; nif3\_YndG: 1.  
 DR PROSITE: PS01136; UPF0034: 1.  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 384 AA; 41100 MW; 44B6D0B1D678E43C CRC64;  
 Query Match 33.3%; Score 5; DB 1; Length 384;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 11 HLRSF 15  
 Db 34 HLRSF 38

RESULT 15  
 GSPF\_AERHY STANDARD; PRT; 388 AA.  
 ID GSPF\_AERHY  
 AC P11743;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE General secretion pathway protein F.  
 GN EXEF.  
 OS Aeromonas hydrophila.  
 OC Bacteria: Proteobacteria: gamma subdivision: Aeromonadaceae;  
 OC Aeromonas.  
 OX NCBI\_TaxID=644;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Ah65;  
 RX MEDLINE=9234963; PubMed=1640836;  
 RA Jiang B., Howard S.P.;  
 RT "The *Aeromonas hydrophila* exef gene, required both for protein  
 RT secretion and normal outer membrane biogenesis, is a member of a  
 RT general secretion pathway.";  
 RL Mol. Microbiol. 6:1351-1361(1992).  
 CC -1- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE  
 CC EXPORT OF PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE PULF/OUTF/EXEF/XPSF/XCPS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X65504; CAA47127.1; -;



Wed Nov 27 08:54:09 2002

DR PIR: S22670; S22670.  
 DR InterPro: IPR003004; Bac\_GSPF.  
 DR InterPro: IPR001992; Bac\_sec\_systII.  
 DR Pfam: PF00462; GSPIL\_F; 1.  
 DR PRINTS: PR00812; BCTERIALGSPF.  
 DR PROSITE: PS00874; T2SP\_F; 1.  
 KW Transport; Transmembrane; Inner membrane.  
 FT TRANSMEM 154 174 POTENTIAL.  
 FT TRANSMEM 206 226 POTENTIAL.  
 FT TRANSMEM 351 371 POTENTIAL.  
 SQ SEQUENCE 388 AA; 43463 MW; D47BI628F977ED50 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRS 14  
 |||||  
 Db 81 AHLRS 85

Search completed: November 27, 2002, 07:28:42  
 Job time : 6.44788 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:24:25 : Search time 4.2278 Seconds  
(Without alignments) 731.044 Million cell updates/sec

Title: US-09-893-615-1

Sequence: 1 WRMYFSHRHAHLRSP 15

Scoring table:

Gapop 60.0, Gapext 60.0,

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	6	40.0	212	16	Q98A19
2	6	40.0	254	17	Q8TU4
3	6	40.0	306	5	Q9BK13
4	6	40.0	352	10	Q23610
5	6	40.0	354	11	Q9CYT9
6	6	40.0	390	4	Q9H6C3
7	6	40.0	390	4	Q96F14
8	6	40.0	449	16	Q53668
9	6	40.0	468	17	Q8TUC6
10	6	40.0	511	2	Q9KH39
11	6	40.0	556	5	Q18509
12	6	40.0	582	2	Q50992
13	6	40.0	633	5	Q9VZK2
14	6	40.0	664	5	Q24140
15	6	40.0	901	5	Q9VA02
16	6	40.0	910	16	Q99U74

17	6	40.0	916	2	Q8VU00	Q8VU00 pseudomonas
18	6	40.0	924	16	Q931R8	Q931R8 staphylococ
19	6	40.0	942	16	Q93J8	Q93J8 neisseria m
20	6	40.0	943	16	Q93J8	Q93J8 pseudomonas
21	6	40.0	943	16	Q93J8	Q93J8 pseudomonas
22	6	40.0	1037	5	Q62495	Q62495 caenorhabd1
23	6	33.3	72	4	Q9H6C3	Q9H6C3 homo sapien
24	6	33.3	82	12	Q8Q017	Q8Q017 camelox vl
25	6	33.3	84	11	Q9R295	Q9R295 mus musculu
26	6	33.3	98	2	Q45250	Q45250 bradyrhizob
27	6	33.3	101	17	Q9VA40	Q9VA40 aeropyrum p
28	6	33.3	119	2	Q9RCE0	Q9RCE0 xanthomonas
29	6	33.3	125	5	Q95T72	Q95T72 drosophila
30	6	33.3	127	4	Q96MT4	Q96MT4 homo sapien
31	6	33.3	142	5	Q20386	Q20386 caenorhabd1
32	6	33.3	142	5	Q95WE4	Q95WE4 phlebotomus
33	6	33.3	142	11	Q9D7M7	Q9D7M7 mus musculu
34	6	33.3	142	11	Q9D1E4	Q9D1E4 mus musculu
35	6	33.3	147	16	Q9H2E1	Q9H2E1 pseudomonas
36	6	33.3	148	16	Q8ZFO4	Q8ZFO4 yersinia pe
37	6	33.3	149	11	Q60874	Q60874 mus musculu
38	6	33.3	151	16	Q8YU06	Q8YU06 anabena sp
39	6	33.3	152	2	Q9AES6	Q9AES6 aeromonas h
40	6	33.3	168	16	Q9S4X5	Q9S4X5 escherichia
41	6	33.3	169	5	Q9VR7	Q9VR7 drosophila
42	6	33.3	170	5	Q9NBF8	Q9NBF8 philodina r
43	6	33.3	176	17	Q54600	Q54600 halobacteri
44	6	33.3	186	5	Q9V69	Q9V69 drosophila
45	6	33.3	190	16	Q53749	Q53749 mycobacteri

## ALIGNMENTS

## RESULT 1

Q98A19 PRELIMINARY: PRT: 212 AA.  
AC Q98A19: 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DR 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein ml16193.  
GN ML16193.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shlimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.,  
RT Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti.  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP003008; BAB52525.1; -  
SO Hypothetical protein: Complete proteome.  
SEQUENCE 212 AA: 2319 MW; 9C52FE82191E38A3 CMC54:

Query Match 40.0%; Score 6; DB 16; Length 212;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 HAHLS 14  
Db 71 HAHLS 76

RESULT 2

08TTU4  
ID 08TTU4 PRELIMINARY: PRT: 254 AA.  
AC 08TTU4:  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Hypothetical protein MA0331.  
GN MA0331.  
OS Methanosarcina acetivorans.  
OC Archaea: Euryarchaeota: Methanococci: Methanosarcinales:  
OC Methanosarcinaceae: Methanosarcina.  
OX NCBI\_Taxid=2214;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN=C2A / ATCC 35395 / DSM 2834;  
RX MEDLINE=21929760; PubMed=11932238;  
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atwood D., Brown A.,  
RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,  
RA Linton L., McMan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,  
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.U., Paulsen I.,  
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
RA Metcalf W.W., Birren B.;  
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
RT and physiological diversity.";  
RL Genome Res. 12:532-542(2002).  
DR EMBL: AE010691; AM03784.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 254 AA; 28615 MW; 627C0A84EA185BF0 CRC64;

Query Match 40.0%; Score 6; DB 17; Length 254;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FSHRHA 10  
111111  
Db 110 FSHRHA 115

RESULT 3  
09BK13  
ID 09BK13 PRELIMINARY: PRT: 306 AA.  
AC 09BK13:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).  
GN GAPC.  
OS Parametium tetraurelia.  
OC Eukaryota: Alveolata: Ciliophora: Oligohymenophorea: Peniculiida:  
OC Parametium.  
OX NCBI\_Taxid=5888;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP MEDLINE=21143069; PubMed=11230543;  
RX Fast N.M., Kissinger J.C., Roos D.S., Keeling P.J.;  
RT "Nuclear-Encoded, Plastid-Targeted Genes Suggest a Single Common  
RT Origin for Apicomplexan and Dinoflagellate Plastids.";  
RL Mol. Biol. Evol. 18:418-426(2001).  
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + PHOSPHATE +  
CC NAD(+) -> 3-PHOSPHO-D-GLYCEROYL PHOSPHATE + NADH.  
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
CC DEHYDROGENASE FAMILY.  
CC EMBL: AF319452; AA020731.1;  
DR HSSP: p56649; IDS.  
DR InterPro: IPR000173; GAP\_dhndrogenase.  
DR Pfam: PF00044; gpoh; 1.  
DR Pfam: PF02800; gpoh\_C; 1.  
DR PRINTS: PR00078; G3PDHDRGNASE.

DR PROSITE: PS00071; GAPDH; 1.  
KW Glycoyls; NAD; Oxidoreductase.  
FT NON\_TER 1  
FT NON\_TER 306  
SQ SEQUENCE 306 AA; 32657 MW; 680D7EC6DDAA527C CRC64;

Query Match 40.0%; Score 6; DB 5; Length 306;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AHLRSP 15  
111111  
Db 41 AHLRSP 46

RESULT 4  
023610  
ID 023610 PRELIMINARY: PRT: 352 AA.  
AC 023610:  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE Hypothetical 40.0 kDa protein.  
GN ATG617700.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;  
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_Taxid=3702;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP MEDLINE=98121113; PubMed=9461215;  
RX Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,  
RA Bergkamp R., Dirkse W., van Staveren M., Slekema W., Drost L.,  
RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Medler H.,  
RA Medler E., Villarroel R., Weitzenger T., Pohl T.M., Terryn N.,  
RA Gielens J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,  
RA Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,  
RA Eutlan K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,  
RA Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,  
RA Vouklatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermayer B.,  
RA Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,  
RA Palmer K., Benes V., Reichman S., Ansoorge W., Cooke R., Berger C.,  
RA Delseny M., Voet M., Voelckert G., Mewes H.W., Klotznerman S.,  
RA Schueller C., Chalmatzis N.;  
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
RT Arabidopsis thaliana.";  
RL Nature 391:485-488(1998).  
RN [2]  
RC SEQUENCE FROM N.A.  
RP EU Arabidopsis sequencing project;  
RX Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL: 297344; CAB10550.1;  
DR EMBL: AL161547; CAB78773.1;  
KW Hypothetical protein.  
SQ SEQUENCE 352 AA; 39986 MW; EF85363502BBE161 CRC64;

Query Match 40.0%; Score 6; DB 10; Length 352;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AHLRSP 15  
111111  
Db 179 AHLRSP 184

RESULT 5  
09CYT9  
ID 09CYT9 PRELIMINARY: PRT: 354 AA.  
AC 09CYT9:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

```

DE 2810450N13R1k protein.
GN 2810450N13R1k
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staibli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection.;
RL Nature 409:685-690(2001).
DR EMBL: AK013322; BAB28790.1;
DR MGD: MGI:1920024; 2810450N13R1k.
DR InterPro: IPR000191; Fapy_DNA_glyco.
DR Pfam: PF01149; Fapy_DNA_glyco; 1.
DR ProDom: PD003680; Fapy_DNA_glyco; 1.
SQ SEQUENCE 354 AA; 39604 MW; 86378B9E94A55CD CRC64;

Query Match 40.0%; Score 6; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RHAHLR 13
DB 95 RHAHLR 100

RESULT 6
ID 09H6C3 PRELIMINARY; PRT; 390 AA.
AC 09H6C3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE CDNA: FLJ22402 fls, clone HRC08117.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuyoshi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isono S., Sugano S.;
RT "NEO human cDNA sequencing project."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AA026055; BAB15337.1;
DR InterPro: IPR000191; Fapy_DNA_glyco.
DR Pfam: PF01149; Fapy_DNA_glyco; 1.
DR ProDom: PD003680; Fapy_DNA_glyco; 1.
SQ SEQUENCE 390 AA; 43685 MW; 8BB040DC8D21B719 CRC64;

Query Match 40.0%; Score 6; DB 4; Length 390;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RHAHLR 13
DB 95 RHAHLR 100

RESULT 7
ID 096F14 PRELIMINARY; PRT; 390 AA.
AC 096F14;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 43.7 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC010876; AAH10876.1;
DR InterPro: IPR000191; Fapy_DNA_glyco.
DR Pfam: PF01149; Fapy_DNA_glyco; 1.
DR ProDom: PD003680; Fapy_DNA_glyco; 1.
SQ SEQUENCE 390 AA; 43712 MW; 2BB040D9EF689382 CRC64;

Query Match 40.0%; Score 6; DB 4; Length 390;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RHAHLR 13
DB 95 RHAHLR 100

RESULT 8
ID 053668 PRELIMINARY; PRT; 449 AA.
AC 053668;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Putative membrane protein.
GN RV0246 OR MT0260 OR MT034.12.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=9825987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLaren J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,

```

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
RA Delcher A., Utterback T., Weldman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL021929; CAAL7338.1; ALT\_INIT.  
DR EMBL: AE006934; AA44478.1; -.  
DR TIGR: MT0260; -.  
DR Tuberculist: RV0246; -.  
KW Complete proteome.  
SO SEQUENCE 449 AA; 46926 MW; FCE6EF930F03D1E6 CRC64;

Query Match 40.0%; Score 6; DB 16; Length 449;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 HAHLRS 14  
DB 2 HAHLRS 7

RESULT 9

Q8TUC6 PRELIMINARY; PRT; 468 AA.  
AC Q8TUC6;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Ethanolamine permease.  
GN MTPP OR MA0143.  
OS Methanosarcina acetivorans.  
OC Archaea: Euryarchaeota: Methanococci: Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2214;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C2A / ATCC 35395 / DSM 2834;  
RX MEDLINE=21929760; PubMed=11932238;  
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
RA Fitzhugh W., Calvo S., Engels R., Smlinov S., Atnoor D., Brown A.,  
RA Allen N., Naylor J., Stange-Thomann N., Dearlano K., Johnson R.,  
RA Litton L., McEwan P., McKernan K., Talamas J., Tlrell A., Ye W.,  
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,  
RA Hedderich R., Ingram-Smith C., Kueltner H.C., Krzycki J.A.,  
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,  
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
RA Metcalf W.W., Birren B.;  
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
RT and physiological diversity.";  
RL Genome Res. 12:532-542(2002).  
DR EMBL: AE010671; AA03556.1; -.  
KW Complete proteome.  
SO SEQUENCE 468 AA; 51099 MW; 3B1D465FE13F77DC CRC64;

Query Match 40.0%; Score 6; DB 17; Length 468;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 YFSHRH 9  
DB 428 YFSHRH 433

RESULT 10

Q9KH39 PRELIMINARY; PRT; 511 AA.  
AC Q9KH39;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE Hypothetical 55.5 kDa protein.  
OS Mycobacterium smegmatis.  
OC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae;  
OC Actinomycetales: Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1772;

RN [1]  
RP SEQUENCE FROM N.A.  
RA Sikder D., Gupta S., Nagaraia V.;  
RT "A hypothetical protein.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF272032; AAF76209.1; -.  
KW Hypothetical protein.  
SO SEQUENCE 511 AA; 55543 MW; 735256A6DE3A035B CRC64;

Query Match 40.0%; Score 6; DB 2; Length 511;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRSP 15  
DB 246 AHLRSP 251

RESULT 11

Q18509 PRELIMINARY; PRT; 556 AA.  
AC Q18509;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE C38D4.4 protein.  
GN C38D4.4.  
OS Caenorhabditis elegans.  
OC Eukaryota: Metazoa: Nematoda: Chromadorea; Rhabditidae; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Coles L.;  
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C. elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL: Z46241; CAA86317.1; -.  
SO SEQUENCE 556 AA; 60548 MW; 79D13559EADBDIAE CRC64;

Query Match 40.0%; Score 6; DB 5; Length 556;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRSP 15  
DB 126 AHLRSP 131

RESULT 12

Q50992 PRELIMINARY; PRT; 582 AA.  
AC Q50992;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE SuCA protein (Fragment).  
GN SUCA.  
OS Neisseria gonorrhoeae.  
OC Bacteria: Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=485;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RC STRAIN-MS11;
RA Porcella S.F., Belland R.J., Judd R.C.;
RT "the sucAB-1pd operon of Neisseria gonorrhoeae.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L36381; AAA96485.1;
DR InterPro: IPR001017; Dehydrogenase_E1.
DR InterPro: IPR004360; Gly_dleo_diox.
DR InterPro: IPR000360; Transketolase.
DR Pfam: PF00676; EL_dehydrog.1.
DR Pfam: PF02779; transket_pyr.1.
DR NON-TR 1
SQ SEQUENCE 582 AA; 65653 MW; 32470A34CF254DEB CRC64;

Query Match 40.0%; Score 6; DB 2; Length 582;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FSHRHA 10
Db 266 FSHRHA 271

RESULT 13
Q9VZK2 PRELIMINARY; PRT; 653 AA.
ID 09VZK2: 09V144;
AC 09VZK2: 09V144;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE SCRT protein (SCRATCH).
GN SCRT OR CG1130.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams W.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anagnostides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adell J.F., Agdayani A., An H.-J., Andrews-Piankovich C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fjosek A., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Foster C., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman J.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jostall M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Kechum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenhach J.,

```

```

RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN 121
RP SEQUENCE OF 1-567 FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agdayani A., Arcalata T.T., Baxter E., Blazet R.G., Butenoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galie R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Paclob J.M.,
RA Park S., Sequeira A., Sethi H., Shtl E., Svirskas R.R., Weinburg T.,
RA Celinker S.E.;
RT "Full length Drosophila melanogaster cDNA sequence.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003479; AAF47819.1;
DR EMBL: AF145627; AAD38602.1;
DR HSSP: P08048; 7ZNF.
DR FLYbase: FBgn0004880; scrt.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; Zf-C2H2; 5.
DR ProDom: PD000003; Znf_C2H2; 1.
DR SMART: SM00355; Znf_C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 3.
DR DNA-binding; Metal-binding; Zinc-finger.
KW SEQUENCE 653 AA; 70774 MW; 15C27EE93FF3F45D CRC64;

Query Match 40.0%; Score 6; DB 5; Length 653;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 HAHURS 14
Db 343 HAHURS 348

RESULT 14
Q24140 PRELIMINARY; PRT; 664 AA.
ID 024140:
AC 024140:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Neuron specific zinc finger transcription factor.
GN SCRT OR SCRATCH OR CG1130.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=96010286; PubMed=7557390;
RA Roark M., Sturtevant M.A., Emery J., Veessin H., Grell E., Bter E.;
RT "scratch, a pan-neural gene encoding a zinc finger protein related to
RT snail, promotes neuronal development.";
RL Genes Dev. 9:2384-2398(1995).
CC -1-SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: U36477; AAA91035.1;
DR HSSP: P08048; 7ZNF.
DR FLYbase: FBgn0004880; scrt.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; Zf-C2H2; 5.
DR PRINTS: PR00048; ZINCFINGER.
DR ProDom: PD000003; Znf_C2H2; 1.
DR SMART: SM00355; Znf_C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 6.
KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
KW SEQUENCE 664 AA; 72569 MW; 687F2DABFE8968F CRC64;

```

Query Match 40.0%; Score 6; DB 5; Length 664;  
 Best Local Similarity 100.0%; Pred No. 46;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 HAHRS 14  
 |||||  
 Db 343 HAHRS 348

## RESULT 15

Q9VA02 PRELIMINARY; PRT; 901 AA.  
 ID Q9VA02  
 AC Q9VA02  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE CG1544 protein.  
 GN CG1544.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003777; AAF57126.1; -;  
 DR FLYBase; FBgn0039827; CG1544.  
 DR InterPro; IPR001017; Dehydrogenase\_E1.  
 DR InterPro; IPR000360; Transketolase.  
 DR Pfam; PF00676; El.dehydrog; 1.  
 DR Pfam; PF02779; transket\_dyr; 1.  
 DR TrnFAMS; TIGR00239; zoxo\_dh\_E1; 1.

SQ SEQUENCE 901 AA; 101506 MM; 656BFCF752E65F83 CRC64;  
 Query Match 40.0%; Score 6; DB 5; Length 901;  
 Best Local Similarity 100.0%; Pred No. 59;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FSHRHA 10  
 |||||  
 Db 584 FSHRHA 589

Search completed: November 27, 2002, 07:30:18  
 Job time : 11.4778 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:19:05 : Search time 5.32819 Seconds  
(Without alignments) 375.129 Million cell updates/sec

Title: US-09-893-615-1

Perfect score: 15  
Sequence: 1 WRMYFSHRHAHLRSP 15

Scoring table: OLIGO  
Gapop-60.0, Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: A\_Geneseq\_101002.\*

```
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	15	100.0	15	20	AAW94701
2	15	100.0	19	20	AAW94733
3	15	100.0	19	20	AAW94714
4	15	100.0	19	20	AAW94715
5	15	100.0	19	20	AAW94717
6	15	100.0	19	20	AAW94718
7	15	100.0	19	20	AAW94719
8	6	40.0	76	23	ABP38325
9	6	40.0	129	22	AAU48733
10	6	40.0	132	23	ABP06906

11	6	40.0	238	22	AAU66845
12	6	40.0	252	20	AAV35921
13	6	40.0	350	20	AAV35922
14	6	40.0	365	22	AAW25386
15	6	40.0	390	20	AAV35923
16	6	40.0	413	20	AAV34811
17	6	40.0	653	22	ABW57870
18	6	40.0	899	22	AAU33860
19	6	40.0	901	22	ABW58224
20	6	40.0	912	22	ABW52560
21	6	40.0	932	20	AAW97701
22	6	40.0	943	22	AAU37021
23	6	40.0	943	22	AAU36241
24	5	33.3	14	22	AAW8006
25	5	33.3	14	22	AAW8007
26	5	33.3	14	22	AAW8008
27	5	33.3	18	22	ABW33959
28	5	33.3	18	22	ABW19398
29	5	33.3	18	22	AAW54723
30	5	33.3	18	22	AAW14984
31	5	33.3	18	22	AAW27417
32	5	33.3	18	22	AAW02709
33	5	33.3	18	23	ABG36782
34	5	33.3	21	23	AAU88558
35	5	33.3	26	21	AAW53840
36	5	33.3	36	20	AAV49674
37	5	33.3	41	23	ABW94072
38	5	33.3	46	22	ABG07585
39	5	33.3	48	22	AAW64327
40	5	33.3	48	22	AAW77154
41	5	33.3	48	22	AAW37285
42	5	33.3	48	23	ABG46166
43	5	33.3	50	22	AAU55721
44	5	33.3	51	22	AAU59920
45	5	33.3	51	22	AAU65435

#### ALIGNMENTS

```
RESULT 1
AAW94701
AAW94701 standard; peptide: 15 AA.
AC AAW94701;
DT 22-APR-1999 (first entry)
XX
XX Lipoteichoic acid epitope peptide mimic for Mab 9C-110.
DE
XX
XX Monoclonal antibody: Mab; lipoteichoic acid; gram positive; bacteria;
KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
KW Mab 9C-110.
XX
XX Staphylococcus sp.
OS
XX
XX W09857994-A2.
XX
XX 23-DEC-1998.
XX
XX 16-JUN-1998; 98WO-US12402.
XX
XX 16-JUN-1997; 97US-0049871.
XX
XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
XX WPI: 1999-095329/08.
XX
XX New antibodies to lipoteichoic acid of gram positive bacteria - used
XX to develop products for the diagnosis, prevention and treatment of
XX infections caused by gram positive bacteria
PT
PT
```

XX Claim 16; Page 120; 150pp; English.  
 PS  
 CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic acid  
 CC of gram positive bacteria, where the Mab is a chimeric immunoglobulin  
 CC comprising at least part of a human immunoglobulin constant region and  
 CC at least part of a non-human immunoglobulin variable region having  
 CC specificity to lipoteichoic acid of gram positive bacteria. The  
 CC antibodies bind to whole bacteria and enhance phagocytosis and killing of  
 CC the bacteria and enhance protection from lethal infection. The antibodies  
 CC or peptides (encoded by a DNA of the variable region of anti-lipoteichoic  
 CC acid antibody or characterised by amino acids corresponding to one or  
 CC more of the complementarily determining regions (CDRs) of the variable  
 CC region of the antibody) can be used for treating or preventing infections  
 CC caused by gram positive bacteria. They can also be used for the diagnosis  
 CC of gram positive bacterial infections. The present sequence represents a  
 CC specifically claimed lipoteichoic acid epitope peptide mimic that can be  
 CC bound by the antibody of the invention (Mab 96-110).  
 CC  
 SQ Sequence 15 AA;

Query Match 100.0%; Score 15; DB 20; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-10;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRWFSHRHAHLRSP 15  
 DB 1 WRWFSHRHAHLRSP 15

RESULT 2  
 AAW94733  
 ID AAW94733 standard; Protein; 19 AA.

AC AAW94733;

DT 22-APR-1999 (first entry)

XX Sequence 15mer 2nd.12 resulting from library panning experiments.

KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KM immunoglobulin; phagocytosis; Infection; epitope; peptide mimic;  
 KM Mab 96-110; panning.

XX Staphylococcus sp.

OS WO9857994-A2.

PN 23-DEC-1998.

PD 16-JUN-1998; 98WO-US12402.

PF 16-JUN-1997; 97US-0049871.

PR (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PA Fischer GW, Schuman RF, Stinson JL, Wong H;

PI MPI: 1999-095329/08.

DR N-PSDB; AAX05559.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 XX

PS Example 5; Fig 8; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarily determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94726-34 represent common-peptide sequences resulting from  
 CC all library panning experiments. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
 CC  
 SQ Sequence 19 AA;

Query Match 100.0%; Score 15; DB 20; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRWFSHRHAHLRSP 15  
 DB 3 WRWFSHRHAHLRSP 17

RESULT 3  
 AAW94714  
 ID AAW94714 standard; Protein; 19 AA.

AC AAW94714;

DT 22-APR-1999 (first entry)

XX Sequence 15mer 2-12/0 resulting from library panning experiment.

KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KM immunoglobulin; phagocytosis; Infection; epitope; peptide mimic;  
 KM Mab 96-110; panning.

XX Staphylococcus sp.

OS WO9857994-A2.

PN 23-DEC-1998.

PD 16-JUN-1998; 98WO-US12402.

PF 16-JUN-1997; 97US-0049871.

PR (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PA Fischer GW, Schuman RF, Stinson JL, Wong H;

PI MPI: 1999-095329/08.

DR N-PSDB; AAX05540.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 XX

PS Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarily determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also

CC be used for the diagnosis of gram positive bacterial infections.  
CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
CC panning second experiment. Three series of panning experiments were  
CC conducted to identify peptide sequences to which antibody of the  
CC invention (Mab 96-110) bound strongly. The translated sequences provide  
CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
XX  
SQ Sequence 19 AA:  
Query Match 100.0%; Score 15; DB 20; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2,2e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WRMYFSHRHAHLRSP 15  
Db 3 WRMYFSHRHAHLRSP 17  
RESULT 4  
AAW94715 standard; Protein: 19 AA.  
XX  
AC AAW94715:  
XX  
DT 22-APR-1999 (first entry)  
XX  
DE Sequence 15mer2-13/0 resulting from 15mer library panning experiment.  
XX  
KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
KW Mab 96-110; panning.  
XX  
OS Staphylococcus sp.  
XX  
PN W09857994-A2.  
XX  
PD 23-DEC-1998.  
XX  
PF 16-JUN-1998; 98WO-US12402.  
XX  
PR 16-JUN-1997; 97US-0049871.  
XX  
PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
PI Fischer GW, Schuman RF, Stinson JL, Wong H;  
XX  
DR WPI: 1999-095329/08.  
DR N-PSDB; AAX05541.  
XX  
PS New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT infections caused by gram positive bacteria  
XX  
XX Example 5; Fig 6A-B; 150pp; English.  
XX  
CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
CC acid of gram positive bacteria, where the Mab is a chimeric  
CC immunoglobulin comprising at least part of a human immunoglobulin  
CC constant region and at least part of a non-human immunoglobulin variable  
CC region having specificity to lipoteichoic acid of gram positive bacteria.  
CC The antibodies bind to whole bacteria and enhance phagocytosis and  
CC killing of the bacteria and enhance protection from lethal infection. The  
CC antibodies or peptides (encoded by a DNA of the variable region of  
CC anti-lipoteichoic acid antibody or characterised by amino acids  
CC corresponding to one or more of the complementarity determining regions  
CC (CDRs) of the variable region of the antibody) can be used for treating  
CC or preventing infections caused by gram positive bacteria. They can also  
CC be used for the diagnosis of gram positive bacterial infections.  
CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
CC panning second experiment. Three series of panning experiments were  
CC conducted to identify peptide sequences to which antibody of the  
CC invention (Mab 96-110) bound strongly. The translated sequences provide  
CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX  
SQ Sequence 19 AA:  
Query Match 100.0%; Score 15; DB 20; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2,2e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WRMYFSHRHAHLRSP 15  
Db 3 WRMYFSHRHAHLRSP 17  
RESULT 5  
AAW94717 standard; Protein: 19 AA.  
XX  
AC AAW94717:  
XX  
DT 22-APR-1999 (first entry)  
XX  
DE Sequence 15mer2-10/0 resulting from 15mer library panning experiment.  
XX  
KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
KW Mab 96-110; panning.  
XX  
OS Staphylococcus sp.  
XX  
PN W09857994-A2.  
XX  
PD 23-DEC-1998.  
XX  
PF 16-JUN-1998; 98WO-US12402.  
XX  
PR 16-JUN-1997; 97US-0049871.  
XX  
PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
PI Fischer GW, Schuman RF, Stinson JL, Wong H;  
XX  
DR WPI: 1999-095329/08.  
DR N-PSDB; AAX05543.  
XX  
PS New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT infections caused by gram positive bacteria  
XX  
XX Example 5; Fig 6A-B; 150pp; English.  
XX  
CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
CC acid of gram positive bacteria, where the Mab is a chimeric  
CC immunoglobulin comprising at least part of a human immunoglobulin  
CC constant region and at least part of a non-human immunoglobulin variable  
CC region having specificity to lipoteichoic acid of gram positive bacteria.  
CC The antibodies bind to whole bacteria and enhance phagocytosis and  
CC killing of the bacteria and enhance protection from lethal infection. The  
CC antibodies or peptides (encoded by a DNA of the variable region of  
CC anti-lipoteichoic acid antibody or characterised by amino acids  
CC corresponding to one or more of the complementarity determining regions  
CC (CDRs) of the variable region of the antibody) can be used for treating  
CC or preventing infections caused by gram positive bacteria. They can also  
CC be used for the diagnosis of gram positive bacterial infections.  
CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
CC panning second experiment. Three series of panning experiments were  
CC conducted to identify peptide sequences to which antibody of the  
CC invention (Mab 96-110) bound strongly. The translated sequences provide  
CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
XX  
SQ Sequence 19 AA:  
Query Match 100.0%; Score 15; DB 20; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2,2e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRMYFSHRAHLRSP 15  
 |||||||||||||  
 DB 3 WRMYFSHRAHLRSP 17

RESULT 6  
 AAW94718  
 ID AAW94718 standard; Protein; 19 AA.  
 AC AAW94718;  
 DT 22-APR-1999 (first entry)  
 DE Sequence 15mer2-16/0 resulting from 15mer library panning experiment.  
 KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KM Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 XX Mab 96-110; panning.  
 OS Staphylococcus sp.  
 PN WO9857994-A2.  
 PD 23-DEC-1998.  
 PF 16-JUN-1998; 98WO-US12402.  
 PR 16-JUN-1997; 97US-0049871.  
 PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
 PI Fischer GW, Schuman RF, Stinson JL, Wong H;  
 XX WPI; 1999-095329/08.  
 DR N-PSDB; AAX05344.  
 PS Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope mimics to which Mab96-110 bound.  
 XX

SQ Sequence 19 AA;  
 Query Match 100.0%; Score 15; DB 20; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRMYFSHRAHLRSP 15  
 |||||||||||||  
 DB 3 WRMYFSHRAHLRSP 17

RESULT 7  
 AAW94719  
 ID AAW94719 standard; Protein; 19 AA.  
 AC AAW94719;  
 DT 22-APR-1999 (first entry)  
 DE Sequence 15mer2-17/0 resulting from 15mer library panning experiment.  
 KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KM Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 XX Mab 96-110; panning.  
 OS Staphylococcus sp.  
 PN WO9857994-A2.  
 PD 23-DEC-1998.  
 PF 16-JUN-1998; 98WO-US12402.  
 PR 16-JUN-1997; 97US-0049871.  
 PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
 PI Fischer GW, Schuman RF, Stinson JL, Wong H;  
 XX WPI; 1999-095329/08.  
 DR N-PSDB; AAX05345.  
 PS Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope mimics to which Mab96-110 bound.  
 XX

SQ Sequence 19 AA;  
 Query Match 100.0%; Score 15; DB 20; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRMYFSHRAHLRSP 15  
 |||||||||||||  
 DB 3 WRMYFSHRAHLRSP 17

RESULT 8  
 ABP38325  
 ID ABP38325 standard; Protein; 76 AA.  
 AC ABP38325;  
 XX

XX

DT 24-JUL-2002 (first entry)  
XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3170.  
DE Staphylococcus epidermidis: open reading frame; ORF; bacterial infection;  
XX Staphylococcus epidermidis: gene therapy.  
KW Staphylococcus epidermidis.  
XX OS  
XX US6380370-B1.  
PN 30-APR-2002.  
XX 13-AUG-1998; 98US-0134001.  
PE 14-AUG-1997; 97US-055779P.  
PR 08-NOV-1997; 97US-064964P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA Doucette-Stamm LA, Bush D;  
PI WPI; 2002-381255/41.  
XX N-PSDB; ABN90870.  
DR  
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
PT polypeptide, useful for diagnosing and treating bacterial infections -  
XX Disclosure: SEQ ID 3170; 267pp; English.  
PS  
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences  
CC can also be used in the diagnosis and treatment of bacterial infections,  
CC particularly S. epidermidis infections. The sequences can be used to  
CC screen for compounds able to interfere with the S. epidermidis life  
CC cycle or inhibit S. epidermidis infection.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site.  
XX  
SQ Sequence 76 AA;  
Query Match 40.0%; Score 6; DB 23; Length 76;  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
GY 5 FSHRHA 10  
DB 10 FSHRHA 15  
RESULT 9  
AAU48733  
ID AAU48733 standard; Protein; 129 AA.  
XX  
XX AAU48733;  
AC  
XX  
XX 27-FEB-2002 (first entry)  
DT  
XX  
XX Propionibacterium acnes immunogenic protein #9629.  
DE  
XX  
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertostis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
XX Propionibacterium acnes.  
OS  
XX  
XX WO200181581-A2.  
PN  
XX  
XX 01-NOV-2001.  
PD

XX  
XX 20-APR-2001; 2001WO-US12865.  
PE  
XX 21-APR-2000; 2000US-199047P.  
PR  
XX 02-JUN-2000; 2000US-208841P.  
PR 07-JUL-2000; 2000US-216747P.  
XX  
XX (CORI-) CORIXA CORP.  
PA Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhattacha A;  
XX L' Maisonneuve J, Zhang Y, Jen S, Carter D;  
PI WPI; 2001-616774/71.  
XX N-PSDB; AAS59543.  
DR  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -  
XX Example 1; SEQ ID NO 9928; 1069pp; English.  
PS  
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertostis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 129 AA;  
Query Match 40.0%; Score 6; DB 22; Length 129;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
GY 10 AHLRSP 15  
DB 78 AHLRSP 83  
RESULT 10  
ABP06906  
ID ABP06906 standard; Protein; 132 AA.  
XX  
XX ABP06906;  
AC  
XX  
XX 24-JUN-2002 (first entry)  
DT  
XX  
XX Human ORF protein sequence SEQ ID NO:13794.  
DE  
XX  
XX Human; open reading frame; ORF; gene therapy; cancer; cleftosis;  
KW hyperproliferative disorder; psoriasis; benign tumor; hemorrhage;  
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myasthenia gravis.  
OS  
XX  
XX Homo sapiens.  
XX

XX WO200192523-A2.  
 XX  
 XX 06-DEC-2001.  
 PD  
 XX  
 XX 29-MAY-2001; 2001WO-US10836.  
 PF  
 XX 30-MAY-2000; 2000US-206132P.  
 PR  
 XX 29-AUG-2000; 2000US-228716P.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Shimkets RA, Leach MD;  
 PI  
 XX WPI; 2002-106308/14.  
 DR  
 XX N-PSDB; ABN22658.  
 DR  
 XX  
 PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders  
 PS  
 XX Disclosure; SEQ ID 13794; 1037pp; English.

XX The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where X is 1-1191 (see Table 1  
 CC in the specification), ABN15762 to ABN27252 encode the human ORFX  
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
 CC treating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage.  
 CC N.B. The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 132 AA;

Query Match 40.0%; Score 6; DB 23; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRHA 10  
 |||||  
 Db 49 FSHRHA 54

RESULT 11  
 AAU66845  
 ID AAU66845 standard; Protein; 238 AA.  
 XX  
 AC AAU66845;

XX 13-FEB-2002 (first entry)  
 XX  
 DE Propionibacterium acnes immunogenic protein #27741.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.  
 XX  
 XX Propionibacterium acnes.  
 OS  
 XX  
 XX WO200181581-A2.  
 PN  
 XX  
 XX 01-NOV-2001.  
 PD  
 XX  
 XX 20-APR-2001; 2001WO-US12865.  
 PF  
 XX 21-APR-2000; 2000US-199047P.  
 PR  
 XX 02-JUN-2000; 2000US-208841P.  
 PR  
 XX 07-JUL-2000; 2000US-216747P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhutia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 DR  
 XX WPI; 2001-616774/71.  
 DR  
 XX N-PSDB; AAS59760.

PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 PS  
 XX Example 1; SEQ ID NO 28040; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 238 AA;

Query Match 40.0%; Score 6; DB 22; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRSP 15  
 |||||  
 Db 187 AHLRSP 192

RESULT 12  
 AAY35921  
 ID AAY35921 standard; Protein; 252 AA.  
 XX  
 AC AAY35921;

XX 13-SEP-1999 (first entry)  
 XX  
 DE Extended human secreted protein sequence, SEQ ID NO. 170.

KW Secreted protein; human; cytokine; cellular proliferation; cell movement;  
 KM cellular differentiation; immune system regulator; anti-inflammatory;  
 KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;

```

KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy.
KM genetic disease.
XX
XX Homo sapiens.
OS
XX WO931236-A2.
PN
XX 24-JUN-1999.
PD
XX
XX 17-DEC-1998; 98WO-IB02122.
PE
XX
XX 10-AUG-1998; 98US-0096116.
PR 17-DEC-1997; 97US-0069957.
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
XX
XX (GEST ) GENSET.
PA
XX
XX Bougueleret L, Duclert A, Dumas Milne Edwards J;
PI
XX
XX WPI; 1999-385906/32.
DR
XX N-PSDB; AAX97605.
XX
XX
XX New isolated human secreted proteins
PT
XX
XX Claim 9; Page 201-202; 516pp; English.
PS
XX
XX This sequence is encoded by an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.
XX
XX
XX Sequence 252 AA:
SQ
XX
XX
XX Query Match 40.0%; Score 6; DB 20; Length 252;
XX Best Local Similarity 100.0%; Pred. No. 31;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 8 RHAHLR 13
XX |||||
DB 95 RHAHLR 100
XX
XX
XX RESULT 13
XX AAY35922
XX AAY35922 standard; Protein; 350 AA.
XX
XX AAY35922;
XX
XX DT 13-SEP-1999 (first entry)
XX
XX Extended human secreted protein sequence. SEQ ID NO. 171.
DE
XX
XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW haemopoietic hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease.
XX
XX Homo sapiens.
OS
XX
XX WO9931236-A2.
PN
XX
XX 24-JUN-1999.
PD

```

```

PF 17-DEC-1998; 98MO-IB02122.
PR XX
PR 10-AUG-1998; 98US-0096116.
PR 17-DEC-1997; 97US-0069957.
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
XX
PA (GEST ) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
XX WPI: 1999-385906/32.
DR N-PSDB: AAX97606.
XX
XX New isolated human secreted proteins
XX
XX Claim 9; Page 203-204; 516pp; English.
XX
XX This sequence is encoded by an extended human secreted protein coding
XX sequence of the invention. The secreted proteins can be used in treating
XX or controlling a variety of human conditions. The secreted proteins may
XX act as cytokines or may affect cellular proliferation or differentiation
XX or may act as immune system regulators, haematopoiesis regulators, tissue
XX growth regulators, regulators of reproductive hormones or cell movement
XX or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
XX tumour inhibition activity. The DNAs can be used in forensic procedures
XX to identify individuals or in diagnostic procedures to identify
XX individuals having genetic diseases resulting from abnormal expression of
XX the genes corresponding to the extended cDNAs. They are also useful for
XX constructing a high resolution map of the human chromosomes. They can
XX also be used for gene therapy to control or treat genetic diseases.
XX
XX Sequence 350 AA;
XX
XX Query Match 40.0%; Score 6; DB 20; Length 350;
XX Best Local Similarity 100.0%; Pred. No. 42;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
OY 8 RHAHR 13
XXXX
DB 95 RHAHR 100
XX
RESULT 14
AAM25386
ID AAM25386 standard; Protein; 365 AA.
XX
XX AAM25386;
XX
XX 16-OCT-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:901.
XX
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX antiinflammatory; antirheumatic; antiarthritis; immunosuppressive;
XX antibacterial; endocrine; cardiac; central nervous system; virucide;
XX anti-HIV; fungicide; antimetagen; cardiovascular; antianemic; anaemia;
XX antagregant; haemostatic; vulnery; antitumor; osteopathic; eczema;
XX dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX cardiac dysfunction; neuropathologic; cardiac anaplylaxis; autoimmune;
XX genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX thrombocytopoenia; osteoporosis; severe combined immunodeficiency;
XX allergic rhinitis; diabetes; multiple sclerosis; depression;
XX Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX neurological disorder.
XX
XX Homo sapiens.
XX
XX WO200153455-A2.
XX

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:26:06 ; Search time 2.08494 Seconds

(without alignments)  
211.682 Million cell updates/sec

Title: US-09-893-615-1

Sequence: 1 WRMYFSHRHAHLRSP 15

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- 1: /cgn2\_6/p10data/1/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/p10data/1/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/p10data/1/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/p10data/1/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/p10data/1/1aa/PCtUS\_COMB.pep.\*
- 6: /cgn2\_6/p10data/1/1aa/backfil1est.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	40.0	76	4 US-09-134-001C-3170	Sequence 3170, Ap
2	6	40.0	932	4 US-08-887-534A-45	Sequence 45, Appl
3	5	33.3	36	4 US-09-183-706-41	Sequence 41, Appl
4	5	33.3	36	4 US-09-567-995-41	Sequence 41, Appl
5	5	33.3	63	2 US-08-338-530A-6	Sequence 6, Appli
6	5	33.3	63	4 US-09-267-384-6	Sequence 6, Appli
7	5	33.3	70	4 US-09-134-001C-3950	Sequence 3950, Ap
8	5	33.3	114	1 US-08-372-952-1	Sequence 1, Appli
9	5	33.3	114	5 US-08-875-309-1	Sequence 1, Appli
10	5	33.3	114	5 PCT-US96-00310-1	Sequence 1, Appli
11	5	33.3	149	2 US-08-815-175-3	Sequence 3, Appli
12	5	33.3	179	3 US-08-665-259-1	Sequence 1, Appli
13	5	33.3	179	3 US-08-762-500-1	Sequence 1, Appli
14	5	33.3	200	4 US-09-562-737-110	Sequence 12, Appl
15	5	33.3	235	1 US-07-940-605A-12	Sequence 12, Appl
16	5	33.3	235	2 US-08-690-096-12	Sequence 12, Appl
17	5	33.3	273	2 US-08-403-853-12	Sequence 12, Appl
18	5	33.3	355	3 US-08-872-979-7	Sequence 7, Appli
19	5	33.3	376	3 US-08-751-512-8	Sequence 8, Appli
20	5	33.3	379	2 US-08-576-626A-50	Sequence 50, Appl
21	5	33.3	385	4 US-09-323-872A-19	Sequence 19, Appl
22	5	33.3	387	4 US-09-916-109-6	Sequence 6, Appli
23	5	33.3	411	2 US-08-338-530A-3	Sequence 3, Appli
24	5	33.3	411	4 US-09-267-384-3	Sequence 3, Appli
25	5	33.3	483	4 US-09-916-109-4	Sequence 4, Appli
26	5	33.3	524	4 US-09-134-001C-4315	Sequence 4315, Ap
27	5	33.3	678	1 US-08-282-141-2	Sequence 2, Appli

28	5	33.3	678	1 US-08-435-434-2	Sequence 2, Appli
29	5	33.3	678	1 US-08-435-436-2	Sequence 2, Appli
30	5	33.3	678	2 US-08-438-863-2	Sequence 2, Appli
31	5	33.3	678	2 US-08-438-864-2	Sequence 2, Appli
32	5	33.3	678	4 US-08-438-862-2	Sequence 2, Appli
33	5	33.3	678	4 US-08-628-747-2	Sequence 2, Appli
34	5	33.3	678	4 US-08-402-253-2	Sequence 2, Appli
35	5	33.3	678	4 US-08-443-866B-2	Sequence 2, Appli
36	5	33.3	840	4 US-08-974-549A-190	Sequence 190, App
37	5	33.3	872	3 US-08-851-843A-8	Sequence 8, Appli
38	5	33.3	872	3 US-08-851-843A-54	Sequence 54, Appli
39	5	33.3	872	4 US-08-974-549A-221	Sequence 221, App
40	5	33.3	872	4 US-08-854-050-8	Sequence 8, Appli
41	5	33.3	872	4 US-08-854-050-54	Sequence 54, Appli
42	5	33.3	872	4 US-09-430-323-8	Sequence 8, Appli
43	5	33.3	872	4 US-09-430-323-54	Sequence 54, Appli
44	5	33.3	979	4 US-08-878-474-5	Sequence 5, Appli
45	5	33.3	1105	4 US-08-999-774A-2	Sequence 2, Appli

## ALIGNMENTS

```
RESULT 1
US-09-134-001C-3170
Sequence 3170, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucelte-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3170
LENGTH: 76
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3170

Query Match          40.0%; Score 6; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 5 FSHRHA 10
Db 10 FSHRHA 15

RESULT 2
US-08-887-534A-45
Sequence 45, Application US/08887534A
Patent No. 6455323
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Horun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,534A  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 28341/33996  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: (312) 474-6600  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 932 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-887-534A-45

Query Match 40.0%; Score 5; DB 4; Length 932;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRHA 10  
|||||  
Db 621 FSHRHA 626

RESULT 3  
US-09-183-706-41  
Sequence 41, Application US/09183706  
Patent No. 6245525  
GENERAL INFORMATION:  
APPLICANT: Martelange, Valrie  
APPLICANT: De Smet, Charles  
TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
FILE REFERENCE: L0461/7054  
CURRENT APPLICATION NUMBER: US/09/183,706  
EARLIER FILING DATE: 1998-10-30  
EARLIER APPLICATION NUMBER: 09/122,989  
EARLIER FILING DATE: 1999-07-27  
NUMBER OF SEQ ID NOS: 43  
SEQ ID NO 41  
LENGTH: 36  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-183-706-41

Query Match 33.3%; Score 5; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRH 9  
|||||  
Db 25 FSHRH 29

RESULT 4  
US-09-567-995-41  
Sequence 41, Application US/09567995  
Patent No. 6303756  
GENERAL INFORMATION:  
APPLICANT: Martelange, Valrie  
APPLICANT: De Smet, Charles  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
FILE REFERENCE: L0461/7054  
CURRENT APPLICATION NUMBER: US/09/567,995  
CURRENT FILING DATE: 2000-05-10  
PRIOR APPLICATION NUMBER: 09/183,706

PRIOR FILING DATE: 1998-10-30  
NUMBER OF SEQ ID NOS: 43  
SEQ ID NO 41  
LENGTH: 36  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-567-995-41

Query Match 33.3%; Score 5; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FSHRH 9  
|||||  
Db 25 FSHRH 29

RESULT 5  
US-08-338-530A-6  
Sequence 6, Application US/08338530A  
Patent No. 5922327  
GENERAL INFORMATION:  
APPLICANT: CRABB, Brendan S.  
APPLICANT: STUDEERT, Michael J.  
TITLE OF INVENTION: EQUINE HERPESVIRUS GLYCOPROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/338,530A  
FILING DATE: 25-JAN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/AU93/00253  
FILING DATE: 28-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40268/120/CSMB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 63 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-338-530A-6

Query Match 33.3%; Score 5; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHRHS 14  
|||||  
Db 43 AHRHS 47

RESULT 6  
US-09-267-384-6  
Sequence 6, Application US/09267384

Patent No. 6193983  
GENERAL INFORMATION:  
APPLICANT: CRABB, Brendan S.  
APPLICANT: STUDDERT, Michael J.  
TITLE OF INVENTION: EQUINE HERPESVIRUS GLYCOPROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/267,384  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/338,530  
FILING DATE: 25-JAN-1995  
APPLICATION NUMBER: WO PCT/AU93/00253  
FILING DATE: 28-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40268/120/CSMB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-3399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 63 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-267-384-6

Query Match 33.3%; Score 5; DB 4; Length 63;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 10 AHLRS 14  
|||||  
Db 43 AHLRS 47

RESULT 7  
US-09-134-001C-3950  
Sequence 3950, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3950  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3950

Query Match 33.3%; Score 5; DB 4; Length 70;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MYFSH 7  
|||||  
Db 52 MYFSH 56

RESULT 8  
US-08-372-952-1  
Sequence 1, Application US/08372952  
Patent No. 5645837  
GENERAL INFORMATION:  
APPLICANT: Jameson, Bradford A.  
APPLICANT: Choksi, Swati  
APPLICANT: Korngold, Robert  
TITLE OF INVENTION: CD8 Antagonists  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESS: No. 564583715  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/372,952  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TDU-1440  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 114 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-372-952-1

Query Match 33.3%; Score 5; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MYFSH 7  
|||||  
Db 102 MYFSH 106

RESULT 9  
US-08-875-309-1  
Sequence 1, Application US/08875309  
Patent No. 6180600  
GENERAL INFORMATION:  
APPLICANT: Jameson, Bradford A.  
APPLICANT: Choksi, Swati  
APPLICANT: Korngold, Robert

APPLICANT: Huang, Ziwei  
TITLE OF INVENTION: CD8 Antagonists  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 61806000115  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,309  
FILING DATE: 26-NOV-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00310  
FILING DATE: 17-JAN-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/372,952  
FILING DATE: 17-JAN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1772  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 114 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-875-309-1

Query Match 33.3%; Score 5; DB 4; Length 114;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MYFSH 7  
|||||  
DB 102 MYFSH 106

RESULT 10  
PCT-US96-00310-1  
Sequence 1, Application PC/TUS9600310  
GENERAL INFORMATION:  
APPLICANT: Jameson, Bradford A.  
APPLICANT: Choksi, Swati  
APPLICANT: Kornogold, Robert  
TITLE OF INVENTION: CD8 Antagonists  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00310  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/372,952  
FILING DATE: 17-JAN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1752  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 114 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-00310-1

Query Match 33.3%; Score 5; DB 5; Length 114;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MYFSH 7  
|||||  
DB 102 MYFSH 106

RESULT 11  
US-08-815-175-3  
Sequence 3, Application US/08815175  
Patent No. 5856139  
GENERAL INFORMATION:  
APPLICANT: Lai, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goll, Surya K.  
TITLE OF INVENTION: NOVEL PROLINE-RICH ACIDIC PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jingle Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/815,175  
FILING DATE: Filed Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0225 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 149 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 899433  
US-08-815-175-3

Query Match 33.3%; Score 5; DB 2; Length 149;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 HLRSR 15  
|||||  
DB 92 HLRSR 96

RESULT 12  
US-08-665-259-1  
; Sequence 1, Application US/08665259  
; Patent No. 6028173  
; GENERAL INFORMATION:  
; APPLICANT: Landes, Gregory M.  
; APPLICANT: Burn, Timothy C.  
; APPLICANT: Connors, Timothy D.  
; APPLICANT: Dackowski, William R.  
; APPLICANT: Van Raay, Terence J.  
; APPLICANT: Klinger, Katherine W.  
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,  
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: One Mountain Road  
; CITY: Framingham  
; STATE: Massachusetts  
; COUNTRY: United States of America  
; ZIP: 01701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/665,259  
; FILING DATE: 17-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dugan, Deborah A.  
; REGISTRATION NUMBER: 37,315  
; REFERENCE/DOCKET NUMBER: IG5-9.1  
; TELEPHONE: (508) 872-8400  
; TELEFAX: (508) 872-5415  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 179 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-665-259-1

Query Match 33.3%; Score 5; DB 3; Length 179;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRS 14  
|||||  
DB 19 AHLRS 23

RESULT 13  
US-08-762-500-1  
; Sequence 1, Application US/08762500

Patent No. 6030806  
; GENERAL INFORMATION:  
; APPLICANT: Landes, Gregory M.  
; APPLICANT: Burn, Timothy C.  
; APPLICANT: Connors, Timothy D.  
; APPLICANT: Dackowski, William R.  
; APPLICANT: Van Raay, Terence J.  
; APPLICANT: Klinger, Katherine W.  
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,  
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME  
; NUMBER OF SEQUENCES: 83  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: One Mountain Road  
; CITY: Framingham  
; STATE: Massachusetts  
; COUNTRY: United States of America  
; ZIP: 01701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/762,500  
; FILING DATE: 09-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/665,259  
; FILING DATE: 17-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10469  
; FILING DATE: 17-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dugan, Deborah A.  
; REGISTRATION NUMBER: 37,315  
; REFERENCE/DOCKET NUMBER: IG5-9.3  
; TELEPHONE: (508) 872-8400  
; TELEFAX: (508) 872-5415  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 179 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-762-500-1

Query Match 33.3%; Score 5; DB 3; Length 179;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 AHLRS 14  
|||||  
DB 19 AHLRS 23

RESULT 14  
US-09-562-737-110  
; Sequence 110, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: URSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 110  
; LENGTH: 200

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-110

```

```

Query Match          33.3%; Score 5; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      6 SHRA 10
        |||||
Db      130 SHRA 134

```

```

RESULT 15
US-07-940-605A-12
; Sequence 12, Application US/07940605A
; Patent No. 5540926

```

```

; GENERAL INFORMATION:
; APPLICANT: ARUFFO, ALEJANDRO
; APPLICANT: HOLLENBAUGH, DIANE
; APPLICANT: LEDBETTER, JEFFREY A.
; TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,605A
; FILING DATE: 04-SEP-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-940-605A-12

```

```

Query Match          33.3%; Score 5; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      3 MYSH 7
        |||||
Db      123 MYSH 127

```

Search completed: November 27, 2002, 07:32:08  
Job time : 4.08494 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:28:06 ; Search time 1.38996 Seconds

(Without alignments)  
171.849 Million cell updates/sec

Title: US-09-893-615-1

Sequence: 1 WRMYFSHRHRLRP 15

Scoring table:

Gap 60.0 , Gapext 60.0

Searched: 102317 seqs, 15924203 residues

Word size : 0

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEM\_PUB pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCRN\_NEM\_PUB pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEM\_PUB pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEM\_PUB pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB pep:\*
- 7: /cgn2\_6/ptodata/1/pubppaa/PCRTUS\_PUBCOMB pep:\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB pep:\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEM\_PUB pep:\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB pep:\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEM\_PUB pep:\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB pep:\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEM\_PUB pep:\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	10	US-09-893-615-1
2	15	100.0	19	10	US-09-893-615-27
3	15	100.0	19	10	US-09-893-615-29
4	15	100.0	19	10	US-09-893-615-33
5	15	100.0	19	10	US-09-893-615-35
6	15	100.0	19	10	US-09-893-615-37
7	15	100.0	19	10	US-09-893-615-65
8	6	40.0	899	10	US-09-815-242-5356
9	6	40.0	932	10	US-09-815-242-12615
10	6	40.0	943	10	US-09-815-242-11834
11	5	33.3	26	10	US-09-864-761-34696
12	5	33.3	26	10	US-09-925-299-1380
13	5	33.3	36	10	US-09-923-831-41
14	5	33.3	48	10	US-09-864-761-45980
15	5	33.3	119	10	US-09-740-288A-10
16	5	33.3	135	10	US-09-949-192-21
17	5	33.3	149	12	US-10-061-690-3
18	5	33.3	254	10	US-09-815-242-5110
19	5	33.3	266	10	US-09-911-826A-6

20	5	33.3	374	10	US-09-740-288A-12	Sequence 12, Appl
21	5	33.3 <td>379</td> <td>10</td> <td>US-09-740-288A-26</td> <td>Sequence 26, Appl</td>	379	10	US-09-740-288A-26	Sequence 26, Appl
22	5	33.3 <td>387</td> <td>10</td> <td>US-09-916-109-6</td> <td>Sequence 6, Appl1</td>	387	10	US-09-916-109-6	Sequence 6, Appl1
23	5	33.3 <td>415</td> <td>10</td> <td>US-09-740-288A-28</td> <td>Sequence 28, Appl1</td>	415	10	US-09-740-288A-28	Sequence 28, Appl1
24	5	33.3 <td>425</td> <td>12</td> <td>US-10-052-586-570</td> <td>Sequence 570, App</td>	425	12	US-10-052-586-570	Sequence 570, App
25	5	33.3 <td>444</td> <td>8</td> <td>US-08-812-393A-2</td> <td>Sequence 2, Appl1</td>	444	8	US-08-812-393A-2	Sequence 2, Appl1
26	5	33.3 <td>460</td> <td>10</td> <td>US-09-925-297-584</td> <td>Sequence 584, App</td>	460	10	US-09-925-297-584	Sequence 584, App
27	5	33.3 <td>467</td> <td>10</td> <td>US-09-870-130-5</td> <td>Sequence 5, Appl1</td>	467	10	US-09-870-130-5	Sequence 5, Appl1
28	5	33.3 <td>469</td> <td>9</td> <td>US-10-001-835-168</td> <td>Sequence 160, App</td>	469	9	US-10-001-835-168	Sequence 160, App
29	5	33.3 <td>483</td> <td>10</td> <td>US-09-916-109-4</td> <td>Sequence 4, Appl1</td>	483	10	US-09-916-109-4	Sequence 4, Appl1
30	5	33.3 <td>500</td> <td>12</td> <td>US-10-006-867-84</td> <td>Sequence 8, Appl1</td>	500	12	US-10-006-867-84	Sequence 8, Appl1
31	5	33.3 <td>872</td> <td>9</td> <td>US-09-843-676-8</td> <td>Sequence 54, Appl1</td>	872	9	US-09-843-676-8	Sequence 54, Appl1
32	5	33.3 <td>872</td> <td>9</td> <td>US-09-843-676-54</td> <td>Sequence 19, Appl</td>	872	9	US-09-843-676-54	Sequence 19, Appl
33	5	33.3 <td>937</td> <td>10</td> <td>US-09-949-192-19</td> <td>Sequence 5, Appl1</td>	937	10	US-09-949-192-19	Sequence 5, Appl1
34	5	33.3 <td>979</td> <td>10</td> <td>US-09-803-187A-5</td> <td>Sequence 5, Appl1</td>	979	10	US-09-803-187A-5	Sequence 5, Appl1
35	5	33.3 <td>1017</td> <td>10</td> <td>US-09-897-056-5</td> <td>Sequence 7, Appl1</td>	1017	10	US-09-897-056-5	Sequence 7, Appl1
36	5	33.3 <td>1025</td> <td>10</td> <td>US-09-897-056-7</td> <td>Sequence 151, App</td>	1025	10	US-09-897-056-7	Sequence 151, App
37	5	33.3 <td>1038</td> <td>10</td> <td>US-09-752-639-151</td> <td>Sequence 151, App</td>	1038	10	US-09-752-639-151	Sequence 151, App
38	5	33.3 <td>1038</td> <td>10</td> <td>US-09-984-198-151</td> <td>Sequence 6, Appl1</td>	1038	10	US-09-984-198-151	Sequence 6, Appl1
39	5	33.3 <td>1275</td> <td>10</td> <td>US-09-749-140-6</td> <td>Sequence 1394, Ap</td>	1275	10	US-09-749-140-6	Sequence 1394, Ap
40	5	33.3 <td>1285</td> <td>10</td> <td>US-09-925-301-1394</td> <td>Sequence 5, Appl1</td>	1285	10	US-09-925-301-1394	Sequence 5, Appl1
41	5	33.3 <td>1474</td> <td>10</td> <td>US-09-873-403-5</td> <td>Sequence 2, Appl1</td>	1474	10	US-09-873-403-5	Sequence 2, Appl1
42	5	33.3 <td>1474</td> <td>12</td> <td>US-10-052-817-2</td> <td>Sequence 54, Appl</td>	1474	12	US-10-052-817-2	Sequence 54, Appl
43	4	26.7	7	9	US-10-080-100-54	Sequence 17, Appl
44	4	26.7	8	9	US-10-080-100-17	Sequence 358, App
45	4	26.7	18	10	US-09-922-261-358	

#### ALIGNMENTS

RESULT 1  
US-09-893-615-1  
Sequence 1, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Mong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPEONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESS: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995, 0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4400  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear

```

; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-893-615-1
Query Match      100.0%; Score 15; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 WRMYFSHRHAHLRSP 15
DB      1 WRMYFSHRHAHLRSP 15

RESULT 2
US-09-893-615-27
; Sequence 27, Application US/09893615
; Patent No. US20020082395A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Gerald W.
;              Schuman, Richard F.
;              Wong, Hing
;              Stinson, Jeffrey L.
; TITLE OF INVENTION: OPOSONIC AND PROTECTIVE MONOCLONAL AND
; CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM
; POSITIVE BACTERIA
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/893,615
; FILING DATE: 29-Jun-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Elnaudt, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 04995.0041-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-893-615-27
Query Match      100.0%; Score 15; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 WRMYFSHRHAHLRSP 15
DB      3 WRMYFSHRHAHLRSP 17

RESULT 3
US-09-893-615-29
; Sequence 29, Application US/09893615
; Patent No. US20020082395A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Gerald W.
;              Schuman, Richard F.
;              Wong, Hing
;              Stinson, Jeffrey L.
; TITLE OF INVENTION: OPOSONIC AND PROTECTIVE MONOCLONAL AND
; CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM
; POSITIVE BACTERIA
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/893,615
; FILING DATE: 29-Jun-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Elnaudt, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 04995.0041-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-893-615-29
Query Match      100.0%; Score 15; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 WRMYFSHRHAHLRSP 15
DB      3 WRMYFSHRHAHLRSP 17

RESULT 4
US-09-893-615-33
; Sequence 33, Application US/09893615
; Patent No. US20020082395A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Gerald W.
;              Schuman, Richard F.
;              Wong, Hing
;              Stinson, Jeffrey L.
; TITLE OF INVENTION: OPOSONIC AND PROTECTIVE MONOCLONAL AND
; CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM
; POSITIVE BACTERIA
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/893,615
; FILING DATE: 29-Jun-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Elnaudt, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 04995.0041-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-893-615-33
Query Match      100.0%; Score 15; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 WRMYFSHRHAHLRSP 15
DB      3 WRMYFSHRHAHLRSP 17
```



SOFTWARE: Patentin Release: #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995,0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-09-893-615-33

Query Match 100.0%; Score 15; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRMYFSHRAHLRSP 15  
|||||  
DB 3 WRMYFSHRAHLRSP 17

RESULT 5  
US-09-893-615-35  
Sequence 35, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995,0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-09-893-615-35

Query Match 100.0%; Score 15; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRMYFSHRAHLRSP 15  
|||||  
DB 3 WRMYFSHRAHLRSP 17

RESULT 6  
US-09-893-615-37  
Sequence 37, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995,0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 37:  
US-09-893-615-37

Query Match 100.0%; Score 15; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRMYFSHRAHLRSP 15  
|||||  
DB 3 WRMYFSHRAHLRSP 17

RESULT 7  
US-09-893-615-65  
Sequence 65, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.

TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-09-893-615-65  
Query Match 100.0%; Score 15; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WRWFSHRHRLSP 15  
DB 3 WRWFSHRHRLSP 17  
RESULT 8  
US-09-815-242-5356  
Sequence 5356, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for windows Version 4.0  
SEQ ID NO 5356  
LENGTH: 899  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5356  
Query Match 40.0%; Score 6; DB 10; Length 899;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FSHRHA 10  
DB 607 FSHRHA 612  
RESULT 9  
US-09-815-242-12615  
Sequence 12615, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for windows Version 4.0  
SEQ ID NO 12615  
LENGTH: 932  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12615  
Query Match 40.0%; Score 6; DB 10; Length 932;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FSHRHA 10  
DB 621 FSHRHA 626  
RESULT 10  
US-09-815-242-11834  
Sequence 11834, Application US/09815242

```
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 11834
LENGTH: 943
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11834

Query Match
Best Local Similarity 40.0%; Score 6; DB 10; Length 943;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FSHRHA 10
|||||
Db 631 FSHRHA 636

RESULT 11
US-09-864-761-34696
Sequence 34696, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34696
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078621.19
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EST_HUMAN HIT: BF059228.1, EVALU02
US-09-864-761-34696

Query Match
Best Local Similarity 33.3%; Score 5; DB 10; Length 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 HLRSP 15
|||||
Db 1 HLRSP 5

RESULT 12
US-09-925-299-1380
Sequence 1380, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1380
LENGTH: 26
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
```

LOCATION: (15)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (19)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (21)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (24)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-923-299-1380

Query Match 33.3%; Score 5; DB 10; Length 26;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 HRRHH 11  
DB 1 HRRHH 5

RESULT 13  
US-09-923-831-41  
Sequence 41, Application US/09923831  
Patent No. US2002015142A1  
GENERAL INFORMATION:  
APPLICANT: Martelange, Val,rie  
APPLICANT: De Smet, Charles  
APPLICANT: Boon-Fallieur, Thierry  
TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
FILE REFERENCE: L0461/7054  
CURRENT APPLICATION NUMBER: US/09/923,831  
CURRENT FILING DATE: 2001-08-07  
PRIOR APPLICATION NUMBER: 09/183,706  
PRIOR FILING DATE: 2001-10-30  
NUMBER OF SEQ ID NOS: 43  
SEQ ID NO 41  
LENGTH: 36  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-923-831-41

Query Match 33.3%; Score 5; DB 10; Length 36;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FSHRH 9  
DB 25 FSHRH 29

RESULT 14  
US-09-864-761-45980  
Sequence 45980, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 45980  
LENGTH: 48  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC010709.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2  
OTHER INFORMATION: EST\_HUMAN HIT: AA852449.1, EVALUOE 5.00e-15  
US-09-864-761-45980

Query Match 33.3%; Score 5; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MYFSH 7  
DB 44 MYFSH 48

RESULT 15  
US-09-740-288A-10  
Sequence 10, Application US/09740288A  
Patent No. US20010039042A1  
GENERAL INFORMATION:  
APPLICANT: Allen, Stephen  
APPLICANT: Kinney, Anthony  
APPLICANT: Miao, Guo-Hua  
APPLICANT: Orozco, Emil  
TITLE OF INVENTION: PLANT BIOTIN SYNTHASE  
FILE REFERENCE: BB1429 US NA  
CURRENT APPLICATION NUMBER: US/09/740,288A  
CURRENT FILING DATE: 2000-12-19  
PRIOR APPLICATION NUMBER: US 60/172929  
PRIOR FILING DATE: 1999-12-21  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 10  
LENGTH: 119

; TYPE: PRT  
; ORGANISM: Argemone mexicana  
US-09-740-288A-10

Query Match 33.3%; Score 5; DB 10; Length 119;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 HRAH 11  
11111  
Db 72 HRAH 76

Search completed: November 27, 2002, 07:32:52  
Job time : 2.38996 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:38:32 : Search time 8.28185 Seconds  
(without alignments)  
241.342 Million cell updates/sec

Title: US-09-893-615-2  
Perfect score: 91  
Sequence: 1 MWMHRIRPLDLAAGR 15

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	91	100.0	15	17	AAW12276
2	91	100.0	15	19	AAW71334
3	91	100.0	15	20	AAW94702
4	91	100.0	15	21	AAW17987
5	91	100.0	15	23	ABW73358
6	91	100.0	19	20	AAW94729
7	91	100.0	19	20	AAW94710
8	91	100.0	19	20	AAW94721
9	91	100.0	19	20	AAW94705
10	91	100.0	19	20	AAW94709

11	91	100.0	37	17	AAW12287
12	80	87.9	19	20	AAW94708
13	69	75.8	15	17	AAW12277
14	56	61.5	11	17	AAW03372
15	56	61.5	15	17	AAW03374
16	51	56.0	159	22	ABG15894
17	50	54.9	11	17	AAW03373
18	50	54.9	15	17	AAW03375
19	50	54.9	139	22	ABB27778
20	50	54.9	139	22	ABB32949
21	50	54.9	139	22	ABB18423
22	50	54.9	139	22	AAW53750
23	50	54.9	139	22	AAW56133
24	50	54.9	139	22	AAW14004
25	50	54.9	139	22	AAW26410
26	50	54.9	139	22	AAW01745
27	50	54.9	139	23	ABG35783
28	49	53.8	157	22	AAW81442
29	49	53.8	157	23	ABW77573
30	46.5	51.1	74	21	AAW54412
31	46	50.5	11	17	AAW03370
32	46	50.5	15	17	AAW03371
33	46	50.5	71	22	AAW06299
34	46	50.5	536	21	AAW84592
35	44.5	48.9	285	23	AAW51648
36	44.5	48.9	298	15	AAW48755
37	44.5	48.9	298	17	AAW07227
38	44.5	48.9	343	23	ABW08348
39	44	48.4	305	22	ABW09714
40	43.5	47.8	74	23	ABP02393
41	43.5	47.8	385	22	AAW93386
42	43	47.3	51	22	AAW56677
43	43	47.3	51	22	AAW29377
44	43	47.3	117	22	AAW64601
45	43	47.3	367	22	ABW11448

#### ALIGNMENTS

RESULT 1	AAW12276	standard; peptide; 15 AA.
ID	AAW12276	
AC	AAW12276;	
DT	15-APR-1997	(first entry)
DE	Synthetic library peptide #1 which binds anti-T. gondii p30 antibody.	
KW	Toxoplasma gondii; surface protein; antibody; screening; peptide library; diagnostic assay; immunisation; phage; fusion protein; envelop.	
OS	Synthetic.	
PN	EP724016-AL.	
PD	31-JUL-1996.	
PF	29-JAN-1996;	96EP-0420030.
PR	30-JAN-1995;	95FR-0001297.
PA	(JOLI/) JOLIVET-REYNARD C.	
PI	(INMR) BIO MERIEUX.	
DR	Jolivet-Reynard C;	
PT	WPI: 1996-343531/35.	
PT	New polypeptide reactive with anti-p30 antibodies against Toxoplasma gondii - useful for diagnosis or immunisation, also new nucleic acid, vectors and transformed cells	

Synthetic template  
Sequence 15mer2-5/  
Synthetic library  
Peptide #6 which b  
Peptide #8 which b  
Novel human diagn  
Peptide #7 which b  
Peptide #9 which b  
Human peptide #429  
Peptide #435 encod  
Protein #422 encod  
Human brain expres  
Human bone marrow  
Peptide #438 encod  
Peptide #447 encod  
Peptide #427 encod  
Human peptide enco  
Human AFP protein  
Human mast cell re  
Zea mays protein f  
Peptide #4 which b  
Peptide #5 which b  
Human foetal prote  
Amino acid sequen  
Rat GPCR polypept  
Human thoracic aor  
RTR-like G protein  
Novel human diagn  
Human ORFX protein  
Human polypeptide,  
Human brain expres  
Peptide #3414 enco  
Protonibacterium  
Drosophila/melanog

XX Example 2; Page 7; 33pp; French.  
PS

The invention relates to novel peptides which bind to antibodies which recognise the *Toxoplasma gondii* P30 envelop protein. A peptide library was generated to express pentadecapeptides on the surface of a filamentous phage as a fusion protein with the pIII protein. The library was screened with immobilised anti-T. *gondii* P30 protein antibody 1E1E7. Phages which bind this antibody were recovered and amplified by one round of infection in *E. coli*. The resultant phages were rescreened with the immobilised antibody and the second round screen isolated 58 bacterial colonies infected with phage. Of the 58 colonies, phage DNA from 30 colonies was isolated and sequenced to determine the sequence of the pentadecapeptide encoded. The peptide sequences AAM12716-66 were identified. Of the 30 colonies studied, this peptide sequence was encoded 11 times. A template peptide sequence (AAM12287) corresponding to the sequence across the phage pIII sequence and putative pentadecapeptide was used to generate a series of overlapping pentadecapeptides. These peptides were used to determine the best peptide sequence which binds the 1E1E7 antibody. Peptides AAM03367-75 were isolated. The new peptides can then be used in diagnostic assays to detect T. *gondii* antibodies in a sample or to purify anti-P30 antibodies or for active immunisation against T. *gondii*.

**SQ Sequence 15 AA;**

Query Match	100.0%;	Score 91;	DB 17;	length 15;
Best Local Similarity	100.0%;	Pred. No. 1.5e-08;		
Matches	15;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

```
QY 1 WHWRHRIPQLAAGR 15
    11111111111111
Db 1 WHWRHRIPQLAAGR 15
```

RESULT 2  
AAW71334  
ID AAW71334 standard; peptide; 15 AA.

AC AA71334;

DT 25-NOV-1998 (first entry)

DE A glycolipid sugar chain peptide.

KW Glycolipid sugar chain; inhiblt; adhesion; metastasis; ~~danger cell.~~

OS Synthetic.

PN JP10237099-A.

PD 08-SEP-1998.

PF 26-FEB-1997; 97JP-0042311.

PR 26-FEB-1997; 97JP-0042311.

PA (IMMO) IMMUNO JAPAN INC.

DR WPI; 1998-537488/46.

PT New peptide which reacts specifically with antibody against  
PT glyco-lipid sugar chains - useful for inhibition of cancer  
PT metastasis

PS Claim 8; page 3; 7pp; Japanese.

CC AAM71332-36 represent glycolipid sugar chain replica peptides. They  
CC react specifically with an antibody against glycolipid sugar chains  
CC and inhibit adhesion and metastasis of cancer cells to a target cell.  
CC The peptides can be used to prevent cancer metastasis.

SD Sequence 15 AA;

Query Match	100.0%	Score 91;	DB 19;	length.15;
Best Local Similarity	100.0%	Pred. No. 1.5e-08;		
Matches	15;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

```

QY      1  WHWHRHRIPLQLAAGR 15
          |||||
DB      1  WHWHRHRIPLQLAAGR 15

```

RESULT 3  
AAW94702  
ID AAW94702 standard; peptide; 15 AA.

AC AAW94702;

DT 22-APR-1999 (first entry)

DE Lipoic acid epitope peptide mimic for Mab 96-110.

KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

KW Mab 96-110.

05 Staphylococcus sp.

PN W09857994-A2.

PD 23-DEC-1998.

~~PF~~ 16-JUN-1998; 98WO-US12402.

PR 16-JUN-1997; 97US-0049871.

PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED

PI Fischer GW, Schuman RF, Stinson JL, Wong H.

DR WPI; 1999-095329/08.

PT New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT infections caused by gram positive bacteria

PS Claim 16; Page 120; 150pp; English.

The invention relates to a monoclonal antibody (Mab) to lipoteichoic acid of gram positive bacteria, where the Mab is a chimeric immunoglobulin comprising at least part of a human immunoglobulin in constant region and at least part of a non-human immunoglobulin variable region having specificity to lipoteichoic acid of gram positive bacteria. The antibodies bind to whole bacteria and enhance phagocytosis and killing of the bacteria and enhance protection from lethal infection. The antibodies or peptides (encoded by a DNA of the variable region of anti-lipoteichoic acid antibody or characterised by amino acids corresponding to one or more of the complementarily determining regions (CDRs) of the variable region of the antibody) can be used for treating or preventing infections caused by gram positive bacteria. They can also be used for the diagnosis of gram positive bacterial infections. The present sequence represents a specifically claimed lipoteichoic acid epitope peptide mimic that can be bound by the antibody of the invention (Mab 96-110).

Sequence	15 AA;
SQ	

Query Match	100.0%;	Score 91;	DB 20;	length 15;
Best Local Similarity	100.0%;	Pred. No. 1.5e-08;		
Matches	15;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

```
QY 1 WHWHRIPQLAAGR 15
    |||||
Db 1 WHWHRIPQLAAGR 15
```



RESULT 4  
 AAB17987  
 ID AAB17987 standard: Peptide: 15 AA.  
 AC AAB17987;  
 DT 31-OCT-2000 (first entry)  
 DE Beta-2bPI Ab binding peptide sequence SEQ ID NO:1099.  
 XX  
 XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
 XX immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;  
 XX MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 XX cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 XX vascular endothelial growth factor; matrix metalloproteinase;  
 XX asthma; thrombosis; pharmaceutical.  
 OS Synthetic.  
 XX  
 XX WO200024782-A2.  
 XX  
 XX PD 04-MAY-2000.  
 XX  
 XX PF 25-OCT-1999; 99WO-US25044.  
 XX  
 XX PR 23-OCT-1998; 98US-0105371.  
 XX PR 22-OCT-1999; 99US-0428082.  
 XX  
 XX (AMGE-) AMGEN INC.  
 XX  
 XX Feige U, Liu C, Cheetham J, Boone TC;  
 XX WPI: 2000-350702/30.  
 XX  
 XX PT Novel composition of matter comprising an Fc domain and  
 XX pharmacologically active peptides, useful for treating cancer and  
 XX autoimmune diseases -  
 XX  
 XX PS Claim 39; Page 598; 608pp; English.  
 XX  
 XX CC The present invention describes composition of matter (I) comprising an  
 XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 XX (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each  
 XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,  
 XX -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4  
 XX where P1, P2, P3, and P4 = are each independently sequences of  
 XX pharmacologically active peptides; L1, L2, L3, and L4 = are each  
 XX independently linkers; and a, b, c, d, e, and f = are each independently  
 XX 0 or 1, provided that at least 1 of a and b is 1. The composition can  
 XX have cytostatic, antiasthmatic, thrombolytic and immunosuppressive  
 XX activities. DNAs, vectors and host cells from the present invention can  
 XX be used for producing pharmaceutical compositions. The compositions are  
 XX useful for treating cancer, asthma, thrombosis, or autoimmune diseases.  
 XX The use of an Fc domain (rather than a Fab domain) can provide a longer  
 XX half-life or incorporate functions such as Fc receptor binding, protein  
 XX A binding, complement fixation, and possibly placental transfer. AAA69443  
 XX to AAA69526 and AAB16953 to AAB18003 represent nucleotide and amino acid  
 XX sequences used in the exemplification of the present invention.  
 XX  
 XX SQ Sequence 15 AA:  
 XX  
 XX Query Match 100.0%; Score 91; DB 21; Length 15;  
 XX Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID ABB73358 standard: Peptide: 15 AA.  
 XX  
 XX ABB73358;  
 XX  
 XX AC  
 XX  
 XX DT 05-APR-2002 (first entry)  
 XX  
 XX DE Exemplary pharmacologically active peptide SEQ ID NO:1057.  
 XX  
 XX XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG;  
 XX EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 XX TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 XX TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 XX MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 XX cytostatic; antineumatic; antiarthritis; antidiabetic; ophthalmological;  
 XX antianemic; anorectic; antiferility; haemostatic; dermatological;  
 XX neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 XX cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 XX sleep disorder; neurological degenerative disease; anaemia;  
 XX thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
 XX Fanconi's syndrome.  
 XX  
 XX OS Synthetic.  
 XX  
 XX XX WO200183525-A2.  
 XX  
 XX XX PD 08-NOV-2001.  
 XX  
 XX PF 02-MAY-2001; 2001WO-US14310.  
 XX  
 XX PR 03-MAY-2000; 2000US-0563286.  
 XX  
 XX (AMGE-) AMGEN INC.  
 XX  
 XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;  
 XX WPI: 2002-130313/17.  
 XX  
 XX PT Novel vehicle-peptide molecule or its multimers useful for treating  
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 XX diabetic retinopathy, obesity, sleep disorders and infertility -  
 XX  
 XX PS Claim 39; Page 62; 176pp; English.  
 XX  
 XX CC The present invention describes a vehicle-peptide molecule (I) or its  
 XX multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 XX cytostatic, antineumatic, antiarthritis, antidiabetic, ophthalmological,  
 XX antianemic, anorectic, antiferility, haemostatic, dermatological and  
 XX neuroprotective activities. (I) can be used as a therapeutic or  
 XX prophylactic agent as well as for screening purposes. (I) is useful for  
 XX diagnosing diseases characterised by dysfunction of their associated  
 XX protein of interest, for identifying normal or abnormal proteins of  
 XX interest, as a part of diagnostic kit to detect the presence of their  
 XX proteins of interest in a biological sample. Additionally, (I) is useful  
 XX for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 XX rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 XX infertility, and neurological degenerative diseases. (I), comprising  
 XX EPO-mimetic compounds are useful for treating disorders characterised by  
 XX low red blood cell levels such as anaemia. The TPO-mimetic comprising  
 XX compounds are useful for treating conditions that involve an existing  
 XX megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 XX deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic  
 XX tumour which result in thrombocytopaenia, systemic lupus erythematosus,  
 XX and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777  
 XX represent amino acid and nucleic acid sequences used in the  
 XX exemplification of the present invention.  
 XX  
 XX SQ Sequence 15 AA:  
 XX  
 XX Query Match 100.0%; Score 91; DB 23; Length 15;  
 XX Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5  
 ABB73358  
 DB 1 WHWHRRIRPIQLAAGR 15

OY 1 WHWHRRIRPIQLAAGR 15

DB 1 WHMRHRIPQLAAGR 15

RESULT 6  
AAW94729

ID AAW94729 standard; Protein; 19 AA.

AC AAW94729;

DT 22-APR-1999 (first entry)

DE Sequence 15mer 2nd.1 resulting from library panning experiments.

KM Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;

KM Mab 96-110; panning.

OS Staphylococcus sp.

PN WO9857994-A2.

PD 23-DEC-1998.

PF 16-JUN-1998; 98WO-US12402.

PR 16-JUN-1997; 97US-0049871.

PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PI Fischer GW, Schuman RF, Stinson JL, Wong H;

DR N-PSDB; AAX05555.

PT New antibodies to lipoteichoic acid of gram positive bacteria - used

PT to develop products for the diagnosis, prevention and treatment of

PT infections caused by gram positive bacteria

PS Example 5; Fig 8; 150pp; English.

CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic

CC acid of gram positive bacteria, where the Mab is a chimeric

CC immunoglobulin comprising at least part of a human immunoglobulin

CC constant region and at least part of a non-human immunoglobulin variable

CC region having specificity to lipoteichoic acid of gram positive bacteria.

CC The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The

CC antibodies or peptides (encoded by a DNA of the variable region of

CC anti-lipoteichoic acid antibody or characterised by amino acids

CC corresponding to one or more of the complementarity determining regions

CC (CDRs) of the variable region of the antibody) can be used for treating

CC or preventing infections caused by gram positive bacteria. They can also

CC be used for the diagnosis of gram positive bacterial infections.

CC Sequences AAW94726-34 represent common peptide sequences resulting from

CC all library panning experiments. Three series of panning experiments were

CC conducted to identify peptide sequences to which antibody of the

CC invention (Mab 96-110) bound strongly. The translated sequences provide

CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX

SQ Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 2e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPQLAAGR 15

DB 3 WHMRHRIPQLAAGR 17

RESULT 7  
AAW94710

ID AAW94710 standard; Protein; 19 AA.

AC AAW94710;

DT 22-APR-1999 (first entry)

DE Sequence 15mer 2-8/0 resulting from 15mer library panning experiment.

KM Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;

KM Mab 96-110; panning.

OS Staphylococcus sp.

PN WO9857994-A2.

PD 23-DEC-1998.

PF 16-JUN-1998; 98WO-US12402.

PR 16-JUN-1997; 97US-0049871.

PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PI Fischer GW, Schuman RF, Stinson JL, Wong H;

DR N-PSDB; AAX05336.

PT New antibodies to lipoteichoic acid of gram positive bacteria - used

PT to develop products for the diagnosis, prevention and treatment of

PT infections caused by gram positive bacteria

PS Example 5; Fig 6A-B; 150pp; English.

CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic

CC acid of gram positive bacteria, where the Mab is a chimeric

CC immunoglobulin comprising at least part of a human immunoglobulin

CC constant region and at least part of a non-human immunoglobulin variable

CC region having specificity to lipoteichoic acid of gram positive bacteria.

CC The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The

CC antibodies or peptides (encoded by a DNA of the variable region of

CC anti-lipoteichoic acid antibody or characterised by amino acids

CC corresponding to one or more of the complementarity determining regions

CC (CDRs) of the variable region of the antibody) can be used for treating

CC or preventing infections caused by gram positive bacteria. They can also

CC be used for the diagnosis of gram positive bacterial infections.

CC Sequences AAW94705-22 represent common peptide sequences resulting from a 15mer library

CC panning second experiment. Three series of panning experiments were

CC conducted to identify peptide sequences to which antibody of the

CC invention (Mab 96-110) bound strongly. The translated sequences provide

CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX

SQ Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 2e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPQLAAGR 15

DB 3 WHMRHRIPQLAAGR 17

RESULT 8  
AAW94721  
ID AAW94721 standard; Protein; 19 AA.  
XX  
AC AAW94721;  
XX  
DT 22-APR-1999 (first entry)

```

DE Sequence 15mer2-19/0 resulting from 15mer library panning experiment.
XX
XX Monoclonal antibody: Mab; lipoteichoic acid; gram positive; bacteria;
XX Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX Mab 96-110; panning.
XX
OS Staphylococcus sp.
XX
XX MO9857994-A2.
XX
XX 23-DEC-1998.
XX
XX 16-JUN-1998; 98MO-US12402.
XX
XX 16-JUN-1997; 97US-0049871.
XX
XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
XX WPI; 1999-095329/08.
XX
XX N-PSDB; AAX05547.
XX
XX
XX New antibodies to lipoteichoic acid of gram positive bacteria - used
XX to develop products for the diagnosis, prevention and treatment of
XX infections caused by gram positive bacteria
XX
XX Example 5; Fig 6A-B; 150pp; English.
XX
XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic
XX acid of gram positive bacteria, where the Mab is a chimeric
XX immunoglobulin comprising at least part of a human immunoglobulin
XX constant region and at least part of a non-human immunoglobulin variable
XX region having specificity to lipoteichoic acid of gram positive bacteria.
XX The antibodies bind to whole bacteria and enhance phagocytosis and
XX killing of the bacteria and enhance protection from lethal infection. The
XX antibodies or peptides (encoded by a DNA of the variable region of
XX anti-lipoteichoic acid antibody or characterised by amino acids
XX corresponding to one or more of the complementarity determining regions
XX (CDRs) of the variable region of the antibody) can be used for treating
XX or preventing infections caused by gram positive bacteria. They can also
XX be used for the diagnosis of gram positive bacterial infections.
XX Sequences AAW94705-22 represent sequences resulting from a 15mer library
XX panning second experiment. Three series of panning experiments were
XX conducted to identify peptide sequences to which antibody of the
XX invention (Mab 96-110) bound strongly. The translated sequences provide
XX lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.
XX
XX Sequence 19 AA:
XX
XX Query Match 100.0%; Score 91; DB 20; Length 19;
XX Best Local Similarity 100.0%; Pred. NO. 2e-08;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
XX QY 1 WHMRRIRPLQLAAGR 15
XX | | | | | | | | | | | | | | |
XX DB 3 WHMRRIRPLQLAAGR 17
XX
XX RESULT 9
XX ID AAW94705
XX AA AAW94705 standard; Protein; 19 AA.
XX
XX AC AAW94705;
XX
XX 22-APR-1999 (first entry)
XX
XX Sequence 15mer2-1/0 resulting from 15mer library panning experiment.
XX
XX Monoclonal antibody: Mab; lipoteichoic acid; gram positive; bacteria;
XX Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX Mab 96-110; panning.
XX

```

```

OS   Staphylococcus sp.
XX
XX   W09857994-A2.
XX
XX   23-DEC-1998.
XX
XX   16-JUN-1998; 98WO-US12402.
XX
XX   16-JUN-1997; 97US-0049871.
XX
XX   (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX   Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
XX   WPI; 1999-095329/08.
XX
XX   N-PSDB; AAX05531.
XX
XX
XX   New antibodies to lipoteichoic acid of gram positive bacteria - used
XX   to develop products for the diagnosis, prevention and treatment of
XX   infections caused by gram positive bacteria
XX
XX   Example 5; Fig 6A-B; 150pp; English.
XX
XX   The invention relates to a monoclonal antibody (Mab) to lipoteichoic
XX   acid of gram positive bacteria, where the Mab is a chimeric
XX   immunoglobulin comprising at least part of a human immunoglobulin
XX   constant region and at least part of a non-human immunoglobulin variable
XX   region having specificity to lipoteichoic acid of gram positive bacteria.
XX   The antibodies bind to whole bacteria and enhance phagocytosis and
XX   killing of the bacteria and enhance protection from lethal infection. The
XX   antibodies or peptides (encoded by a DNA of the variable region of
XX   anti-lipoteichoic acid antibody or characterised by amino acids
XX   (CPRS) of the variable region of the antibody) can be used for treating
XX   or preventing infections caused by gram positive bacteria. They can also
XX   be used for the diagnosis of gram positive bacterial infections.
XX   Sequences AAW94705-22 represent sequences resulting from a 15mer library
XX   panning second experiment. Three series of panning experiments were
XX   conducted to identify peptide sequences to which antibody of the
XX   invention (Mab 96-110) bound strongly. The translated sequences provide
XX   lipoteichoic acid epitope peptide mimics to which mab96-110 bound.
XX
XX   Sequence 19 AA:
XX
XX   Query Match 100.0%; Score 91; DB 20; Length 19;
XX   Best Local Similarity 100.0%; Pred. No. 2e-08;
XX   Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX   QY 1 WHMRRRIPLQLAAGR 15
XX   |||||
XX   DB 3 WHMRRRIPLQLAAGR 17
XX
XX
XX   RESULT 10
XX   ID AAW94709 standard; Protein; 19 AA.
XX
XX   AAW94709;
XX
XX   22-APR-1999 (first entry)
XX
XX   Sequence 15mer2-7/0 resulting from 15mer library panning experiment.
XX
XX   Monoclonal antibody: Mab; lipoteichoic acid; gram positive; bacteria;
XX   Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX   Mab 96-110; panning.
XX
XX   Staphylococcus sp.
XX
XX   W09857994-A2.
XX
XX   23-DEC-1998.
XX

```

```

PI Jolivet-Reynaud C;
XX
DR WPI; 1996-343531/35.
XX
PT New polypeptide reactive with anti-P30 antibodies against Toxoplasma
XX gondii - useful for diagnosis or immunisation, also new nucleic
XX acid, vectors and transformed cells
XX
PS Example 5; Page 9; 33pp; French.
XX
CC The invention relates to novel peptides which bind to antibodies which
CC recognise the Toxoplasma gondii P30 envelop protein. A peptide library
CC was generated to express pentadecapeptides on the surface of a
CC filamentous phage as a fusion protein with the P11 protein. The library
CC was screened with immobilised anti-T. gondii P30 protein antibody 1E1E7.
CC Phages which bind this antibody were recovered and amplified by one
CC round of infection in E. coli. The resultant phages were rescreened with
CC the immobilised antibody and the second round screen isolated 58
CC bacterial colonies infected with phage. Of the 58 colonies, phage DNA
CC from 30 colonies was isolated and sequenced to determine the sequence of
CC the pentadecapeptide encoded. The peptide sequences AAW12276-86 were
CC identified. Of the 30 colonies studied, this peptide sequence was
CC encoded once. A template peptide sequence (AAW12287) corresponding to
CC the sequence across the phage P11 sequence (AAW12287) corresponding to
CC pentadecapeptide was used to generate a series of overlapping
CC pentadecapeptides. These peptides were used to determine the best
CC peptide sequence which binds the 1E1E7 antibody. Peptides AAW03367-75
CC were isolated. The new peptides can then be used in diagnostic assays to
CC detect T. gondii antibodies in a sample or to purify anti-P30 antibodies
CC or for active immunisation against T. gondii.
XX
SQ Sequence 37 AA;
XX
Query Match 100.0%; Score 91; DB 17; Length 37;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 1 WWHRRRRIPLQLAAGR 15
XXXXXXXXXXXXXXXXXXXX
DB 10 WWHRRRRIPLQLAAGR 24
XXXXXXXXXXXXXXXXXXXX
RESULT 12
AAW94708
ID AAW94708 standard; Protein; 19 AA.
XX
AC AAW94708;
XX
DT 22-APR-1999 (first entry)
XX
DE Sequence 15mer2-5/0 resulting from 15mer library panning experiment.
XX
KM Monoclonal antibody; Mab; Iipoteichoic acid; gram positive; bacteria;
XX immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX Mab 96-110; panning.
XX
KM Staphylococcus sp.
XX
OS
XX
PN WO9857994-A2.
XX
PD 23-DEC-1998.
XX
PF 16-JUN-1998; 98WO-US12402.
XX
PR 16-JUN-1997; 97US-0049871.
XX
PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
PI Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
DR WPI; 1999-095329/08.
XX
DR N-PSDB; AAX05534.
XX

```

PT New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 XX  
 XX Example 5; Fig 6A-B; 150pp; English.  
 PS  
 CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
 CC  
 XX Sequence 19 AA:  
 SO  
 OY Query Match 87.9%; Score 80; DB 20; Length 19;  
 Db Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 HMRHRIPQLAAGR 15  
 Db 4 HMRHRIPQLAAGR 17  
 ID AAW12277 standard; peptide: 15 AA.  
 AC AAW12277;  
 XX  
 DT 15-APR-1997 (first entry)  
 XX  
 DE Synthetic library peptide #2 which binds anti-T. gondii P30 antibody.  
 XX  
 KW Toxoplasma gondii; surface protein; antibody; screening; peptide library;  
 KW diagnostic assay; immunisation; phage; fusion protein; envelop.  
 XX  
 OS Synthetic.  
 XX  
 PN EP724016-A1.  
 PD 31-JUL-1996.  
 XX  
 PF 29-JAN-1996; 96EP-0420030.  
 XX  
 PR 30-JAN-1995; 95FR-0001297.  
 PA (JOLI/) JOLIVET-REYNAUD C.  
 PA (INMR) BIO MERIEUX.  
 XX  
 PI Jolivet-Reynaud C;  
 XX  
 DR WPI; 1996-343531/35.  
 XX  
 PT New polypeptide reactive with anti-P30 antibodies against Toxoplasma  
 PT gondii - useful for diagnosis or immunisation, also new nucleic  
 PT acid, vectors and transformed cells  
 XX  
 PS Example 2; Page 7; 33pp; French.  
 CC The invention relates to novel peptides which bind to antibodies which

CC recognise the Toxoplasma gondii P30 envelop protein. A peptide library  
 CC was generated to express pentadecapeptides on the surface of a  
 CC filamentous phage as a fusion protein with the PIII protein. The library  
 CC was screened with immobilised anti-T. gondii P30 protein antibody 1E1E7.  
 CC Phages which bind this antibody were recovered and amplified by one  
 CC round of infection in E. coli. The resultant phages were rescreened with  
 CC the immobilised antibody and the second round screen isolated 58  
 CC bacterial colonies infected with phage. Of the 58 colonies, phage DNA  
 CC from 30 colonies was isolated and sequenced to determine the sequence of  
 CC the pentadecapeptide encoded. The peptide sequences AAW12276-86 were  
 CC identified. Of the 30 colonies studied, this peptide sequence was  
 CC encoded once. A template peptide sequence (AAW12287) corresponding to  
 CC the sequence across the phage PIII sequence and putative  
 CC pentadecapeptide was used to generate a series of overlapping  
 CC pentadecapeptides. These peptides were used to determine the best  
 CC peptide sequence which binds the 1E1E7 antibody. Peptides AAW03367-75  
 CC were isolated. The new peptides can then be used in diagnostic assays to  
 CC detect T. gondii antibodies in a sample or to purify anti-P30 antibodies  
 CC or for active immunisation against T. gondii.  
 XX  
 SO Sequence 15 AA:  
 OY Query Match 75.8%; Score 69; DB 17; Length 15;  
 Db Best Local Similarity 85.7%; Pred. No. 7.2e-05;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 2 HMRHRIPQLAAGR 15  
 Db 2 HMRHRIPQLAAGR 15  
 ID AAW03372 standard; peptide: 11 AA.  
 AC AAW03372;  
 XX  
 DT 15-APR-1997 (first entry)  
 XX  
 DE Peptide #6 which binds anti-Toxoplasma gondii P30 antibody.  
 XX  
 KW Toxoplasma gondii; surface protein; antibody; screening; peptide library;  
 KW diagnostic assay; immunisation; phage; E.coli.  
 XX  
 OS Synthetic.  
 XX  
 PN EP724016-A1.  
 PD 31-JUL-1996.  
 XX  
 PF 29-JAN-1996; 96EP-0420030.  
 XX  
 PR 30-JAN-1995; 95FR-0001297.  
 PA (JOLI/) JOLIVET-REYNAUD C.  
 PA (INMR) BIO MERIEUX.  
 XX  
 PI Jolivet-Reynaud C;  
 XX  
 DR WPI; 1996-343531/35.  
 XX  
 DR N-PSDB; AAT31328.  
 XX  
 PT New polypeptide reactive with anti-P30 antibodies against Toxoplasma  
 PT gondii - useful for diagnosis or immunisation, also new nucleic  
 PT acid, vectors and transformed cells  
 XX  
 PS Claim 7; Page 19; 33pp; French.  
 XX  
 CC Peptides AAW03367-75 are peptides isolated from a peptide library, which  
 CC bind to antibody 1E1E7 which recognises the P30 surface protein from  
 CC Toxoplasma gondii. The peptide library was constructed by inserting a  
 CC nucleic acid sequence encoding a random pentadecapeptide in the PIII  
 CC envelop protein of a filamentous phage. Antibody 1E1E7 was immobilised

CC on the bottom of a Petri dish and the expressed peptide library was  
CC overlaid on the antibody. After washing, phages bound on the antibodies  
CC were recovered and the phage amplified by passage through *E. coli*. The  
CC phages were rescreened by the same method and 58 bacterial colonies were  
CC recovered. Of these, 30 were further studied and the nucleic acid  
CC encoding the pentadecapeptide was isolated. The nucleic acids encoded  
CC the peptides AAW12276-86. A series of overlapping peptides based on the  
CC peptide template AAW12287 was constructed to identify which peptide  
CC sequences bound to the anti-P30 antibody. The peptide scanning isolated  
CC peptides AAW03367-75. The peptides can then be used in diagnostic  
CC assays to detect *T. gondii* antibodies in a sample or to purify anti-P30  
CC antibodies or for active immunisation against *T. gondii*.

XX Sequence 11 AA:

Query Match 61.5%; Score 56; DB 17; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0078;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 HRIPQLAAGR 15  
DB 1 HRIPQLAAGR 11

RESULT 15  
AAW03374

ID AAW03374 standard; peptide; 15 AA.

XX AAW03374;

DT 15-APR-1997 (first entry)

DE Peptide #8 which binds anti-Toxoplasma gondii P30 antibody.

KM Toxoplasma gondii: surface protein; antibody: screening; peptide library;  
KM diagnostic assay; immunisation; phage; *E. coli*.

XX Synthetic.

PN EP724016-A1.

PD 31-JUL-1996.

PF 29-JAN-1996; 96EP-0420030.

PR 30-JAN-1995; 95FR-0001297.

PA (JOLI/) JOLIVET-REYNAUD C.

PA (INMR ) BIO MERIEUX.

XX Jolivet-Reynaud C;

PI WPI: 1996-343531/35.

DR N-PSDB; AAT131330.

XX New polypeptide reactive with anti-P30 antibodies against Toxoplasma  
PT gondii - useful for diagnosis or immunisation, also new nucleic  
PT acid, vectors and transformed cells

PS Claim 7; Page 20; 33pp: French.

XX Peptides AAW03367-75 are peptides isolated from a peptide library, which  
CC bind to antibody 1E17 which recognises the P30 surface protein from  
CC Toxoplasma gondii. The peptide library was constructed by inserting a  
CC nucleic acid sequence encoding a random pentadecapeptide in the pIII  
CC envelope protein of a filamentous phage. Antibody 1E17 was immobilised  
CC on the bottom of a Petri dish and the expressed peptide library was  
CC overlaid on the antibody. After washing, phages bound on the antibodies  
CC were recovered and the phage amplified by passage through *E. coli*. The  
CC phages were rescreened by the same method and 58 bacterial colonies were  
CC recovered. Of these, 30 were further studied and the nucleic acid  
CC encoding the pentadecapeptide was isolated. The nucleic acids encoded  
CC the peptides AAW12276-86. A series of overlapping peptides based on the

CC peptide template AAW12287 was constructed to identify which peptide  
CC sequences bound to the anti-P30 antibody. The peptide scanning isolated  
CC peptides AAW03367-75. The peptides can then be used in diagnostic  
CC assays to detect *T. gondii* antibodies in a sample or to purify anti-P30  
CC antibodies or for active immunisation against *T. gondii*.

XX Sequence 15 AA:

Query Match 61.5%; Score 56; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 HRIPQLAAGR 15  
DB 1 HRIPQLAAGR 11

Search completed: November 27, 2002, 07:17:50  
Job time : 11.2819 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:15:05 ; Search time 4.1197 Seconds  
(without alignments)  
350.688 Million cell updates/sec

Title: US-09-893-615-2

Perfect score: 91

Sequence: 1 WHMRHRIQLAAGR 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	51.6	649	2	C75112 hypothetical prote
2	46.5	51.1	354	2	T36559 probable epoxide h
3	46.5	51.1	391	2	T13182 integrase - Lactob
4	46	50.5	96	4	ODEC31 hypothetical prote
5	45	49.5	516	2	H87369 tryptophan halogen
6	44.5	48.9	270	2	C83167 probable transcrip
7	44.5	48.9	343	2	A35639 G protein-coupled
8	44.5	48.9	356	2	F70636 probable epnb prot
9	44	48.4	636	2	D82679 hypothetical prote
10	44	48.4	1839	1	KRWPEM genome polyprotein
11	43	47.3	152	2	G81175 darp pyrophosphohy
12	43	47.3	152	2	H81930 probable nucleosid
13	43	47.3	222	2	T35942 probable ABC-type
14	43	47.3	330	2	G87401 epoxide hydrolase
15	43	47.3	434	2	E82162 paraquat-inducible
16	43	47.3	615	2	T37981 ABC transpoxyst c
17	43	47.3	716	2	AC2449 ABC transporter AT
18	43	47.3	1298	2	A48999 protein-lysine k
19	42	46.2	207	2	A75475 probable acetyltra
20	42	46.2	289	2	T08176 glucose-1-phosphat
21	42	46.2	410	2	T50718 hypothetical prote
22	42	46.2	436	2	B70510 probable acyltrans
23	42	46.2	437	1	E64553 conserved hypothet
24	42	46.2	641	2	C84726 probable receptor-
25	42	46.2	782	2	S62583 probable transcrip
26	42	46.2	1101	2	AB2481 hypothetical prote
27	42	46.2	1189	2	D75516 acetyl-CoA synthas
28	41.5	45.6	642	2	D75516 acetyl-CoA synthas
29	41	45.1	158	2	S39754 pathogenesis-relat

30	41	45.1	232	2	S58353	CD1b protein - she
31	41	45.1	249	2	E87597	transcription regu
32	41	45.1	291	2	C71362	hypothetical prote
33	41	45.1	322	2	B70957	probable epna prot
34	41	45.1	333	2	S47246	gene C01 protein -
35	41	45.1	354	2	JC4212	nitrilase (EC 3.5.
36	41	45.1	389	1	S36653	kpsd protein - Esc
37	41	45.1	435	2	F82888	GTP-binding protei
38	41	45.1	502	2	E87596	tryptophan halogen
39	41	45.1	503	2	H87438	tryptophan halogen
40	41	45.1	509	2	G87596	tryptophan halogen
41	41	45.1	580	2	T36393	L-aspartate oxidas
42	41	45.1	650	2	D71021	hypothetical prote
43	41	45.1	716	2	AH2319	glycyl-tRNA synth
44	41	45.1	1227	2	AG2504	hypothetical prote
45	40.5	44.5	321	2	C84664	epoxide hydrolase

#### ALIGNMENTS

##### RESULT 1

C75112 hypothetical protein PAB1903 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: C75112

R:Anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome s

A:Reference number: A75001

A:Accession: C75112

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-649 <KAW>

A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49612.1; PID:g545

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1903

C:Superfamily: hypothetical protein YPI009c

Query Match 51.6%; Score 47; DB 2; Length 649;

Best Local Similarity 61.5%; Pred. No. 11;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WHMRHRIQLAAG 14

DB 563 NMHGIPLKLAAG 575

##### RESULT 2

T36559 probable epoxide hydrolase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000

C:Accession: T36559

R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A

submitted to the EMBL Data Library, June 1999

A:Reference number: Z21584

A:Accession: T36559

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-354 <SEP>

A:Cross-references: EMBL:AL079353; PIDN:CAB45554.1; GSPDB:GN00070; SCOEDB:SC17.08c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC17.08c

C:Superfamily: troplnesterase

Query Match 51.1%; Score 46.5; DB 2; Length 354;

Best Local Similarity 53.3%; Pred. No. 7.3;

Matches 8; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 1 WH-WRRIPQLAAG 14  
 |||:|:|  
 Db 62 WYSNRHQLPALAAG 76

## RESULT 3

T13182

Integrase - Lactobacillus phage phi-gle

C:Species: Lactobacillus phage phi-gle

C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 21-Jul-2000

C:Accession: T13182

R:Kobata, K.I.; Oki, M.; Kakikawa, M.; Watanabe, N.; Hirakawa, M.; Yamada, K.; Takeo, Gene 187, 45-53, 1997

A:Title: Genome structure of the Lactobacillus temperate phage phi gle: the whole genome

A:Reference number: 217631; MUID:97225795; PMID:9073065

A:Accession: T13182

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-391 &lt;KOD&gt;

A:Cross-references: EMBL:X98106; NID:91926320; PIDN:CAA6758.1; PID:91926326

C:Genetics:

A:Gene: Int

C:Superfamily: phage L54a excisionase

Query Match

Best Local Similarity 51.1%; Score 46.5; DB 2; Length 391;

Matches 8; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

OY 1 WHWR-----HRIPLAAG 14

Db 12 WHWRNRTIDGQRPVINSAG 32

## RESULT 4

O0EC31

hypothetical protein C-96 rspsa 3'-region - Escherichia coli (fragment)

C:Species: Escherichia coli

C:Date: 29-Sep-1999 #sequence\_revision 29-Sep-1999 #text\_change 29-Sep-1999

C:Accession: A04448

R:Schneider, J.; Isono, K.

Nucleic Acids Res. 10, 1857-1865, 1982

A:Title: The DNA sequence of the gene rspsa of Escherichia coli coding for ribosomal prot

A:Reference number: A04448; MUID:62196865; PMID:6261725

A:Accession: A04448

A:Molecule type: DNA

A:Residues: 1-96 &lt;SCH&gt;

A:Cross-references: GB:V00352; GB:J01681; GB:J01682; NID:942899

C:Comment: This is the hypothetical translation of a sequence that was not reported as a

C:Genetics:

A:Map position: 20 min

Query Match

Best Local Similarity 50.5%; Score 46; DB 4; Length 96;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 WHWRHRIPL 9

Db 1 WHWRHRIPL 9

## RESULT 5

H87369

tryptophan halogenase, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001

C:Accession: H87369

R:Niernann, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Debby, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: H87369

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-516 <STO>  
 A:Cross-references: GB:AE005673; NID:913422252; PIDN:AAK22956.1; GSPDB:GN00148  
 C:Genetics:

C:Superfamily: Rhizobium plasmid PNGR234a y4xg protein

Query Match

Best Local Similarity 49.5%; Score 45; DB 2; Length 516;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 WRHRIPLAAG 14

Db 272 WRHRIPLAAG 283

## RESULT 6

C83167

probable transcription regulator PA3830 [imported] - Pseudomonas aeruginosa (strain P

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: C83167

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Micooguchi, S.D.; Warren, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10584043

A:Accession: C83167

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-270 &lt;STO&gt;

A:Cross-references: GB:AE004800; GB:AE004091; NID:9949999; PIDN:AA607217.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3830

Query Match

Best Local Similarity 48.9%; Score 44.5; DB 2; Length 270;

Matches 9; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

OY 2 WHWRHRIPL-----QLAAG 14

Db 213 WHWRHRIPLAALPRLAAG 230

## RESULT 7

A35639

G protein-coupled receptor RTA - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 16-Jul-1999

C:Accession: A35639

R:Ross, P.C.; Figler, R.A.; Corjay, M.H.; Barber, C.M.; Adam, N.; Marcus, D.R.; Lynch

Proc. Natl. Acad. Sci. U.S.A. 87, 3052-3056, 1990

A:Title: RTA, a candidate G protein-coupled receptor: cloning, sequencing, and tissue

A:Reference number: A35639; MUID:90222168; PMID:2105324

A:Accession: A35639

A:Molecule type: mRNA

A:Residues: 1-343 &lt;ROS&gt;

A:Cross-references: GB:M35297; NID:9206809; PIDN:AAA2087.1; PID:9206810; GB:M32098

C:Superfamily: mas transforming protein

C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

F;4/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 48.9%; Score 44.5; DB 2; Length 343;

Matches 8; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

OY 1 WHWRHRIPLAAG 14

Db 150 WHWRHRIPLAAG 162



## RESULT 8

F70636

probable ephB protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: F70636

R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70636

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-356 &lt;COL&gt;

A:Cross-references: GB:284498; GB:AL123456; NID:93261701; PIDN:CAB06523.1; PID:91806224

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: ephB

C:Superfamily: triphosphatase

Query Match                      48.9%    Score 44.5; DB 2; Length 356;  
 Best Local Similarity    53.3%    Pred. No. 15;  
 Matches    8; Conservative    2; Mismatches    4; Indels    1; Gaps    1;

Oy    1 WH-WRRHRIPLQIAG 14  
 1: ||||| 1: 1  
 Db    40 WYSMRHRIPLAGAG 54

## RESULT 9

DB2679

hypothetical protein XF1451 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: DB2679

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: AB2515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number AS9328 below

A:Accession: DB2679

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-636 &lt;SIM&gt;

A:Cross-references: GB:AE003975; GB:AE003849; NID:99106468; PIDN:AAF84260.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Bioness, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Melo, E.; Docena, C.; El-Dorry, H.; Facincanli, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miranda, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J.; de M. de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: AS9328

A:Contents: annotation

C:Genetics:

A:Gene: XF1451

Query Match                      48.4%    Score 44; DB 2; Length 636;  
 Best Local Similarity    60.0%    Pred. No. 33;  
 Matches    6; Conservative    2; Mismatches    2; Indels    0; Gaps    0;

Oy    1 WHWRRHRIPLQ 10  
 1: ||| 1: 1

Db    207 WRRHSHVPLQ 216

## RESULT 10

RWMPEN

genome polyprotein - eggplant mosaic virus

N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase

N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: eggplant mosaic virus

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 19-Jan-2001

C:Accession: J00102

R:Osorio-Keese, M.E.; Keese, P.; Glbbs, A.

Virology 172, 547-554, 1989

A:Title: Nucleotide sequence of the genome of eggplant mosaic tymovirus.

A:Reference number: J00102; MUID:90021185; PMID:2800336

A:Accession: J00102

A:Molecule type: genomic RNA

A:Residues: 1-1839 &lt;OSO&gt;

A:Cross-references: EMBL:J04374

C:Superfamily: eggplant mosaic virus RNA-directed RNA polymerase

C:Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis

F:965-972/Region: nucleotide-binding motif A (P-loop)

F:1027-1032/Region: nucleotide-binding motif B

F:971/Binding site: ATP (Lys) #status predicted

Query Match                      48.4%    Score 44; DB 1; Length 1839;  
 Best Local Similarity    63.6%    Pred. No. 93;  
 Matches    7; Conservative    2; Mismatches    2; Indels    0; Gaps    0;

Oy    3 WRRHRIPLQIAG 13  
 1: |||| 1: 1  
 Db    1094 WRRHRIPLQIAG 1104

## RESULT 11

GB1175

dATP pyrophosphohydrolase NMB0642 [imported] - Neisseria meningitidis (strain MC58 se

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001

C:Accession: GB1175

R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

rl, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: AB1000; MUID:20175755; PMID:10710307

A:Accession: GB1175

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-152 &lt;TEP&gt;

A:Cross-references: GB:AE002419; GB:AE002098; NID:97223563; PIDN:AAF41065.1; PID:9722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0642

Query Match                      47.3%    Score 43; DB 2; Length 152;  
 Best Local Similarity    85.7%    Pred. No. 11;  
 Matches    6; Conservative    0; Mismatches    1; Indels    0; Gaps    0;

Oy    2 HWRHRIPLQ 8  
 1: |||| 1: 1  
 Db    84 HWRHRIPLQ 90

## RESULT 12

HB1930

probable nucleoside triphosphate pyrophosphohydrolase NMA0852 [imported] - Neisseria

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001

C:Accession: HB1930

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jags, J.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.  
A:Reference number: A81775; MUID:20222556; PMID:10761919  
A:Accession: H81930  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <PAR>  
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84133.1; PID:g737956  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: nlpA; NMA0852

Query Match 47.3%; Score 43; DB 2; Length 152;  
Best Local Similarity 85.7%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HMRHRIP 8  
|||||  
Db 84 HMRHRYP 90

RESULT 13  
T35942  
Probable ABC-type transport system ATP-binding protein - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Jan-2000  
C:Accession: T35942  
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Randeram, M.A.;  
submitted to the EMBL Data Library, January 1999  
A:Reference number: Z21551  
A:Accession: T35942  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-222 <SEE>  
A:Cross-references: EMBL:AL035206; PIDN:CAA22764.1; GSPDB:GN00070; SCORDB:SC9B5.22  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCORDB:SC9B5.22  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 47.3%; Score 43; DB 2; Length 222;  
Best Local Similarity 85.7%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HMRHRIP 8  
|||||  
Db 68 HMRHRAP 74

RESULT 14  
G87401  
epoxide hydrolase [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: G87401  
R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.;  
B.; Laub, M.T.; Debby, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: G87401  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-330 <STO>  
A:Cross-references: GB:AE005673; NID:g13422557; PIDN:AAK23211.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CCI229  
C:Superfamily: trophoblastase

Query Match 47.3%; Score 43; DB 2; Length 330;  
Best Local Similarity 61.5%; Pred. No. 25;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 WMRHRIPLOAGR 15  
|||||  
Db 50 WMRHRIPLOAGR 62

RESULT 15  
E82162  
paragut-inducible protein A VC1753 [imported] - *Vibrio cholerae* (strain N16961 serog  
C:Species: *Vibrio cholerae*  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: E82162  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.  
charson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: E82162  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-434 <HEI>  
A:Cross-references: GB:AE004252; GB:AE003852; NID:g9656263; PIDN:AAF94903.1; GSPDB:GN  
C:Experimental source: serogroup O1, strain N16961; biotype El Tor  
A:Gene: VC1753  
A:Map position: 1  
C:Superfamily: hypothetical protein H1671

Query Match 47.3%; Score 43; DB 2; Length 434;  
Best Local Similarity 50.0%; Pred. No. 32;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 WMRHRIPLOAGR 14  
|||||  
Db 204 WMRHRIPLOAGR 217

Search completed: November 27, 2002, 07:24:13  
Job time: 10.112 secs





RESULT 4  
VGR3\_HUMAN STANDARD: PRT: 1298 AA.  
AC P35916;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Vascular endothelial growth factor receptor 3 precursor (EC 2.7.1.112)  
DE (VEGFR-3) (Tyrosine-protein kinase receptor FLT4).  
GN FLT4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Placenta;  
RX MEDLINE=93241723; PubMed=8386825;  
RA Galland F., Karaymshava A., Pebusque M.-J., Borg J.-P., Rotlapel R.,  
RA Dubreuil P., Rosnet O., Birnbaum D.;  
RT "The FLT4 gene encodes a transmembrane tyrosine kinase related to the  
RT vascular endothelial growth factor receptor.";  
RT Oncogene 8:1233-1240(1993).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93007958; PubMed=1327515;  
RA Pajusola K., Aprelikova O., Korhonen J., Kaipainen A., Pertovaara L.,  
RA Alltalo R., Alltalo K.;  
RT "FLT4 receptor tyrosine kinase contains seven immunoglobulin-like  
RT loops and is expressed in multiple human tissues and cell lines.";  
RL Cancer Res. 52:5738-5743(1992).  
[3]  
RP ERRATUM.  
RA Pajusola K., Aprelikova O., Korhonen J., Kaipainen A., Pertovaara L.,  
RA Alltalo R., Alltalo K.;  
RL Cancer Res. 53:3845-3845(1993).  
[4]  
RP SEQUENCE FROM N.A.  
RA Lee J., Gray A., Yuan J., Luch S.M., Avraham H., Wood W.I.;  
RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.  
[5]  
RP SEQUENCE OF 761-1190 FROM N.A.  
RX MEDLINE=92119639; PubMed=1310071;  
RA Aprelikova O., Pajusola K., Partanen J., Armstrong E., Alltalo R.,  
RA Bailey S.K., McMahon J., Masumoto J., Huebner K., Alltalo K.;  
RT "FLT4 a novel class III receptor tyrosine kinase in chromosome 5q33-  
RT qter.";  
RL Cancer Res. 52:746-748(1992).  
CC -1- FUNCTION: RECEPTOR FOR VEGF-C. HAS A TYROSINE-PROTEIN KINASE  
CC ACTIVITY.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: PLACENTA, LUNG, HEART, AND KIDNEY. DOES NOT  
CC SEEM TO BE EXPRESSED IN PANCREAS AND BRAIN.  
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X69878; CAA49505.1; -;  
DR EMBL: X68203; CAA48290.1; ALT\_INIT.  
DR EMBL: U63143; AAB5215.1; -;  
DR PIR: S36130; S36130.  
DR HSP: P11362; 1PGR.  
DR Genew; HGNC:3767; FLT4.

DR MM: 136352; -;  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003598; Ig\_C2.  
DR InterPro: IPR003600; Ig\_Like.  
DR InterPro: IPR001824; RTK\_inseII.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00047; Ig\_5.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 2.  
DR SMART: SM00410; Ig\_Like; 2.  
DR SMART: SM00408; Igc2; 2.  
DR SMART: SM00219; Tyrc; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_LIT; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase: Tyrosine-protein kinase. Phosphorylation; ATP-binding;  
KW Receptor; Transmembrane; Signal; Immunoglobulin domain; Repeat;  
KW Glycoprotein.  
FT SIGNAL 1 23  
FT CHAIN 24 1298  
FT FT  
FT DOMAIN 24 775  
FT TRANSMEM 776 797  
FT DOMAIN 798 1298  
FT DOMAIN 44 118  
FT DOMAIN 151 213  
FT DOMAIN 245 317  
FT DOMAIN 351 403  
FT DOMAIN 438 541  
FT DOMAIN 571 660  
FT DOMAIN 692 758  
FT DOMAIN 845 1173  
FT NP\_BIND 851 859  
FT BINDING 879 879  
FT ACT\_SITE 1037 1037  
FT DISULFID 51 111  
FT DISULFID 158 206  
FT DISULFID 252 310  
FT DISULFID 445 534  
FT DISULFID 578 653  
FT DISULFID 699 751  
FT CARBOHYD 33 33  
FT CARBOHYD 104 104  
FT CARBOHYD 166 166  
FT CARBOHYD 251 251  
FT CARBOHYD 299 299  
FT CARBOHYD 411 411  
FT CARBOHYD 515 515  
FT CARBOHYD 527 527  
FT CARBOHYD 564 564  
FT CARBOHYD 683 683  
FT CARBOHYD 690 690  
FT CARBOHYD 758 758  
FT MOD\_RES 1068 1068  
FT CONFLICT 24 24  
FT CONFLICT 745 745  
FT CONFLICT 752 753  
FT CONFLICT 890 890  
FT CONFLICT 1128 1128  
FT CONFLICT 1146 1146  
FT CONFLICT 1164 1164  
SQ SEQUENCE 1298 AA; 145598 MW; 3DC469ED3C8B83B1 CRC64;  
Query Match 47.3%; Score 43; DB 1; Length 1298;  
Best local Similarity 46.7%; Pred. No. 29;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
OY 1 WHRRRRIPLQLAAGR 15  
DB 458 WHRRPWTCKMFAQR 472

```

RESULT 5
Y269_HELPY
ID Y269_HELPY STANDARD: PRT: 437 AA.
AC P56131:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein HP0269.
GN HP0269.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
ON NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khakh H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Colton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0004 FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch.)
CC -----
CC DR EMBL: AE000546; AAD07337.1; -
CC DR TIGR: HP0269; -
CC DR InterPro: IPR001861; UPF0004.
CC DR Pfam: PF00919; UPF0004. 1.
CC DR TIGRFAMs: TIGR00089; UPF0004. 1.
CC DR PROSITE: PS01278; UPF0004. 1.
CC KM Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 437 AA; 49423 MW; 3ECD4656C1CEB3F CRC64;

Query Match 46.2%; Score 42; DB 1; Length 437;
Best Local Similarity 38.5%; Pred. No. 14;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 3 WRRHPIQLAAGR 15
   |::|::|::|
Db 341 WKERVPLEVSSSR 353

RESULT 6
YAKB_SCHPO
ID YAKB_SCHPO STANDARD: PRT: 782 AA.
AC 009922;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative transcriptional regulatory protein C1F7.11c.
GN SPAC1F7.11c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Nibbelk D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Trivey A., Walsh S.V., Warren F., Whitehead S.,
RA Woodward J., Voickeys G., Aert R., Robben J., Gymnopoulos B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Filiz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Egger P., Zimmermann W., Medler H., Wambuit R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallade V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spharakski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC CLUSTER DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch.)
CC -----
CC DR EMBL: Z67998; CA91958.1; -
CC DR HSSP: P12351; 1HWT.
CC DR InterPro: IPR001138; Fungal_TRN.
CC DR Pfam: PF00172; zn.clus. 1.
CC DR SMART: SM00066; GAL4. 1.
CC DR PROSITE: PS00463; ZN2_CY6_FUNGAL_1; 1.
CC DR PROSITE: PS50048; ZN2_CY6_FUNGAL_2; 1.
CC KM Hypothetical protein; Transcription regulation; DNA-binding;
CC KW Nuclear protein; Zinc; Metal-binding.
CC FT DNA_BIND 22..50 ZN(2)-CYS(6), FUNGAL-TYPE.
CC SQ SEQUENCE 782 AA; 89377 MW; B08DAEC986B81C7F CRC64;

Query Match 46.2%; Score 42; DB 1; Length 782;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWRHRIP 8
   |||
Db 189 HWSHRIP 195

RESULT 7
PRL_ASPOF
ID PRL_ASPOF STANDARD: PRT: 158 AA.
AC 005736;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pathogenesis-related protein 1 (AOPR1).
GN PRL.
OS Asparagus officinalis (Garden asparagus).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC   Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Asparagusceae;
CC   Asparagus.
OX   NCBI_TaxID=4686;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92329716; PubMed=1627770;
RA   Warner S.A.J., Scott R., Draper J.;
RT   "Characterisation of a wound-induced transcript from the monocot
RT   asparagus that shares similarity with a class of intracellular
RT   pathogenesis-related (PR) proteins.";
RL   Plant Mol. Biol. 19:555-561(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94035138; PubMed=8220442;
RA   Warner S.A.J., Scott R., Draper J.;
RT   "Isolation of an asparagus intracellular PR gene (AoPRI) wound-
RT   responsive promoter by the inverse polymerase chain reaction and its
RT   characterization in transgenic tobacco.";
RL   Plant J. 3:191-201(1993).
CC   -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -|- INDUCTION: BY WOUNDING AND ELICITOR TREATMENTS.
CC   -|- SIMILARITY: BELONGS TO THE BETV1 FAMILY OF PATHOGENESIS-RELATED
CC   PROTEINS.
-----
CC   CC
CC   The SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
-----
CC   DR   EMBL; X62103; CAA44013.1; -
DR   EMBL; X64452; CAA45784.1; -
DR   EMBL; A26571; CAA01827.1; -
DR   PIR; S22533; S22533.
DR   PIR; S39754; S39754.
DR   HSSP; O24248; IE09.
DR   InterPro: IPR000916; Bet_v-I.
DR   Pfam: PF00407; Bet_v-I_1.
DR   PRINTS; PR00634; BETALLERGEN.
DR   PROSITE; PS00451; PATHOGENESIS_BETV1; 1.
KW   Pathogenesis-related protein.
SQ   SEQUENCE 158 AA; 16921 MW; C9B9354812B00C20 CRC64;

Query Match          45.1%; Score 41; DB 1; Length 158;
Best Local Similarity 46.2%; Pred. No. 7.5;
Matches    6; Conservative    3; Mismatches      4; Indels     0; Gaps     0;

OY       3 WRHRIPLQLAGR 15
           | : : : ||||
DB        6 WSHEVAVNVNAGR 18

RESULT 8
CIB3_SHEEP
AC   P80943;
ID   CIB3_SHEEP         STANDARD;             PRT; 232 AA.
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   T-cell surface glycoprotein CD1b-3 (CD1b-3 antigen) (SCDIT10)
DE   (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal thymocytes;
RX   MEDLINE=96269982; PubMed=8662069;
RA   Ferguson E.E., Dutla B.M., Hein W.R., Hopkins J.;

```

```

CC "The shsp CD1 gene family contains at least four CD18 homologues." ;
CC RL Immunogenetics 44:66-96(1996).
CC CC
CC -1- FUNCTION: NOT KNOWN.
CC -1- SUBUNIT: ASSOCIATES NON-COVALENTLY WITH BETA-2-MICROGLOBULIN (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO THE IMMUGLOBULIN SUPERFAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage: by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: X90567; CAA62187.1; -.
CC DR HSSP: P11609; ICD1.
CC DR InterPro: IPR003106; Ig_MHC.
CC DR InterPro: IPR003597; Ig_CL.
CC DR Pfam: PF00047; Ig; 1.
CC DR SMART: SM00407; IgC1; 1.
CC KW Glycoprotein; Immunoglobulin domain; Transmembrane; Multigene family.
CC FT NON_TER
CC FT DOMAIN 1 201 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 202 222 POTENTIAL.
CC FT DOWN 223 232 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 19 83 BY SIMILARITY.
CC PF DISULFID 123 178 BY SIMILARITY.
CC FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 232 AA; 26023 MW; C96DB93840B56158 CRC64;
CC
CC Query Match 45.1%; Score 41; DB 1; Length 232;
CC Best Local Similarity 50.0%; Pred. NO. 11;
CC Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
CC
CC Oy 1 WHRRHRIPLOLAAG 14
CC | | | | | | | | | |
CC Db 158 WTWYLRVTLDDVAAG 171
CC
CC RESULT 9
CC Y126-TREPA STANDARD: PRT: 291 AA.
CC AC 083163;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein TP0126.
CC GN TP0126.
CC OS Treponema pallidum.
CC CC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
CC OX NCBI_TaxID=160;
CC [1]
CC SEQUENCE FROM N.A.
CC RP STRAIN=Nichols;
CC RC MEDLINE=98332770; PubMed=9665876;
CC RX Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
CC RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
CC RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
CC RA Khakh H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
CC RA McDonald L., Artlach P., Bowman G., Cotton M.D., Fujii C., Garland S.,
CC RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
CC RA Venter J.C.;
CC RT "Complete genome sequence of Treponema pallidum, the syphilis
CC spirochete." ;
CC RL Science 281:375-388(1998).
CC CC -1- SIMILARITY: SOME, TO T.PALLIDUM TP0733.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE001198; AAC65118.1; -.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 74 96 POTENTIAL.  
 FT TRANSMEM 188 210 POTENTIAL.  
 SO SEQUENCE 291 AA; 31921 MW; 496381BD06E67A3 CRC64;  
 Query Match 45.18; Score 41; DB 1; Length 291;  
 Best Local Similarity 58.38; Pred. No. 14;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Oy 3 WHRRIPLOLAG 14  
 Db 206 WRRLPLVAG 217  
 RESULT 10  
 C1B2\_SHEEP STANDARD; PRT; 333 AA.  
 AC Q29422;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE T-cell surface glycoprotein CD1b-2 precursor (CD1b-2 antigen)  
 DE (SCD1b-42) (Antigen IAH-CC14);  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 NC NCBI\_TaxID:9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymus;  
 RX MEDLINE=96269982; Pubmed=8662069;  
 RA Ferguson E.E., Dutla B.M., Hein W.R., Hopkins J.;  
 RT "The sheep CD1 gene family contains at least four CD1B homologues";  
 RL Immunogenetics 44:86-96(1996).  
 RN [2]  
 RP SEQUENCE OF 21-33.  
 RX MEDLINE=99115506; Pubmed=9914336;  
 RA Rind S.M., Hopkins J., Dutla B.M.;  
 RT "Amino-terminal sequencing of sheep CD1 antigens and identification of  
 a sheep CD1 gene";  
 RL Immunogenetics 49:225-230(1999).  
 CC -1- FUNCTION: NOT KNOWN.  
 CC -1- SUBUNIT: ASSOCIATES NON-COVALENTLY WITH BETA-2-MICROGLOBULIN (BY  
 SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z36891; CAA85360.1; -.  
 DR HSSP: P11609; ICD1.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003597; Ig\_c1.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00407; IgC1; 1.  
 KW Glycoprotein; Signal; Immunoglobulin domain; Transmembrane;  
 KM Multigene family.  
 FT SIGNAL 1 20  
 FT CHAIN 21 333 T-CELL SURFACE GLYCOPROTEIN CD1B-2.  
 FT DOMAIN 21 302 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 303 323 POTENTIAL.  
 FT DOMAIN 324 333 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 120 184 BY SIMILARITY.  
 FT DISULFID 224 279 BY SIMILARITY.  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 333 AA; 37039 MW; 861BAE9617DB9BA1 CRC64;  
 Query Match 45.18; Score 41; DB 1; Length 333;  
 Best Local Similarity 50.08; Pred. No. 16;  
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Oy 1 WHRRIPLOLAG 14  
 Db 259 WTWRLRVTLNVAAG 272  
 RESULT 11  
 KSS5\_ECOLI STANDARD; PRT; 389 AA.  
 AC P42218;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Capsule polysaccharide export protein kpsS.  
 GN KPS.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K5;  
 RX MEDLINE=9388530; Pubmed=8397187;  
 RA Pazzanti C., Rosenow C., Boulnois G.J., Bronner D., Jann K.;  
 RA Roberts I.S.;  
 RT "Molecular analysis of region 1 of the Escherichia coli K5 antigen  
 gene cluster: a region encoding proteins involved in cell surface  
 RT expression of capsular polysaccharide";  
 RL J. Bacteriol. 175:5978-5983(1993).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X74567; CAA52659.1; -.  
 DR Polysaccharide transport; Transport.  
 KW SEQUENCE 389 AA; 46381 MW; 73058122C28027DE CRC64;  
 Query Match 45.18; Score 41; DB 1; Length 389;  
 Best Local Similarity 62.38; Pred. No. 18;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 WHRRIP 8  
 Db 174 WHRRHEP 181  
 RESULT 12  
 PSAB\_GUTH STANDARD; PRT; 734 AA.  
 AC O78507;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Photosystem I P700 chlorophyll A apoprotein A2 (psab) (PSI-B).  
 GN PSAB.  
 OS Guillardia theta (Cryptomonas phl).



OG Chloroplast.  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Gulliardia.  
 OK NCBI\_TaxID=55329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99128221; PubMed=9929392;  
 RA Douglas S.E., Penny S.L.;  
 RT "The plastid genome of the cryptophyte alga, Gulliardia theta:  
 RT complete sequence and conserved synteny groups confirm its common  
 RT ancestry with red algae."  
 RL J. Mol. Evol. 48:236-244(1999).  
 CC -1- FUNCTION: Psaa and psab bind P700, the primary electron donor of  
 CC photosystem I (PSI), as well as the electron acceptors A0, A1, and  
 CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin  
 CC oxidoreductase.  
 CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1  
 CC is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.  
 CC -1- SUBUNIT: A psaa/B heterodimer binds the P700 chlorophyll special  
 CC pair and subsequent electron acceptors. The PSI reaction center of  
 CC higher plants and algae is composed of one at least 11 subunits.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast  
 CC thylakoid membrane.  
 CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF041468; AAC35698.1; -  
 DR HSSP: P25897; IJBO.  
 DR InterPro: IPR001280; PSI\_Psaa/B.  
 DR Pfam: PF00223; psaa\_psab; 1.  
 DR PRINTS: PR00257; PHOTOSPSAB.  
 DR PROSITE: PS00419; PHOTOSYSTEM\_I\_PSAAB; 1.  
 KW Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;  
 KW Iron-sulfur; 4Fe-4S; Chlorophyll.  
 FT TRANSMEM 46 69 I (POTENTIAL).  
 FT TRANSMEM 135 158 II (POTENTIAL).  
 FT TRANSMEM 175 199 III (POTENTIAL).  
 FT TRANSMEM 273 291 IV (POTENTIAL).  
 FT TRANSMEM 330 353 V (POTENTIAL).  
 FT TRANSMEM 369 395 VI (POTENTIAL).  
 FT TRANSMEM 417 439 VII (POTENTIAL).  
 FT TRANSMEM 517 535 VIII (POTENTIAL).  
 FT TRANSMEM 575 596 IX (POTENTIAL).  
 FT TRANSMEM 643 665 X (POTENTIAL).  
 FT TRANSMEM 707 727 XI (POTENTIAL).  
 FT METAL 559 568 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).  
 FT METAL 568 559 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).  
 FT BINDING 654 654 P700 SPECIAL PAIR CHLOROPHYLL AXIAL  
 FT BINDING 654 654 LIGAND (BY SIMILARITY).  
 FT BINDING 662 662 A0 CHLOROPHYLL (BY SIMILARITY).  
 FT BINDING 670 670 A1 CHLOROPHYLL (BY SIMILARITY).  
 FT BINDING 671 671 A1 PHYTYLQUINONE (BY SIMILARITY).  
 FT BINDING 700 700 A1 PHYTYLQUINONE (BY SIMILARITY).  
 SQ SEQUENCE 734 AA; 82328 MW; 823FEAA29E159BD4 CRC64;

Query Match 45.1%; Score 41; DB 1; Length 734;  
 Best Local Similarity 46.7%; Pred. No. 35;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WHMRHRIPLAAGR 15  
 DB 594 WHMRHRIPLAAGR 608

RESULT 13  
 DYHG\_CHLRE

ID DYHG\_CHLRE STANDARD; PRT; 4485 AA.  
 AC Q39575;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Dynein gamma chain, flagellar outer arm.  
 GN ODA2 OR ODA-2.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OK NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=1132D;  
 RX MEDLINE=94274766; PubMed=7516341;  
 RA Wilkerson C.G., King S.M., Wilman G.B.;  
 RT "Molecular analysis of the gamma heavy chain of Chlamydomonas  
 RT flagellar outer-arm dynein".  
 RL J. Cell Sci. 107:497-506(1994).  
 CC -1- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND  
 CC FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.  
 CC DYNEIN HAS ATPASE ACTIVITY.  
 CC -1- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND  
 CC GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.  
 CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U15303; AAA50455.1; -  
 DR InterPro: IPR004273; Dynein\_heavy.  
 DR InterPro: IPR001230; Preyln\_site.  
 DR Pfam: PR03028; Dynein\_heavy; 1.  
 KW Motor protein; Microtubules; Dynein; ATP-binding; Flagella;  
 KW Coiled coil.  
 FT DOMAIN 449 469 COILED COIL (POTENTIAL).  
 FT DOMAIN 804 838 COILED COIL (POTENTIAL).  
 FT DOMAIN 1093 1114 COILED COIL (POTENTIAL).  
 FT DOMAIN 1275 1297 COILED COIL (POTENTIAL).  
 FT DOMAIN 1699 1727 COILED COIL (POTENTIAL).  
 FT DOMAIN 1917 1945 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 3077 3099 COILED COIL (POTENTIAL).  
 FT DOMAIN 3196 3227 COILED COIL (POTENTIAL).  
 FT DOMAIN 3265 3343 COILED COIL (POTENTIAL).  
 FT DOMAIN 3569 3663 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1819 1826 ATP (POTENTIAL).  
 FT NP\_BIND 2099 2106 ATP (POTENTIAL).  
 FT NP\_BIND 2425 2432 ATP (POTENTIAL).  
 FT NP\_BIND 2802 2809 ATP (POTENTIAL).  
 SQ SEQUENCE 4485 AA; 512836 MW; 974B79328B403677 CRC64;

Query Match 44.5%; Score 40.5; DB 1; Length 4485;  
 Best Local Similarity 60.0%; Pred. No. 2,5e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 W-HMRHRCVPV 2382  
 DB 2373 W-HMRHRCVPV 2382

RESULT 14  
 LEGT\_HUMAN STANDARD; PRT; 135 AA.  
 ID LEGT\_HUMAN  
 AC P47929;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Galectin-7 (HKL-14) (P17).

GN LGALST.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 7-19; 75-82; 111-117 AND 120-132.  
 RC TISSUE-Epidermis; PubMed=7534301;  
 RX MEDLINE=95197604; PubMed=7534301;  
 RA Madson P., Rasmussen H.H., Flint T., Gromov P., Kruse T.A., Honore B.,  
 RA Vorum H., Cells J.E.;  
 RA "Cloning, expression, and chromosome mapping of human galectin-7.";  
 RL J. Biol. Chem. 270:5823-5829(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Epidermis;  
 RX MEDLINE=95246905; PubMed=7729568;  
 RA Magnaldo T., Bernard F., Darmon M.;  
 RA "Galectin-7, a human 14-kDa S-lectin, specifically expressed in  
 RT keratinocytes and sensitive to retinoic acid.";  
 RL Dev. Biol. 168:259-271(1995).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=98434364; PubMed=9760227;  
 RA Leonidas D.D., Vatzaki E.H., Vorum H., Cells J.E., Madson P.,  
 RA Acharya K.R.;  
 RA "Structural basis for the recognition of carbohydrates by human  
 RT galectin-7."; Biochemistry 37:13930-13940(1998).  
 CC -1- FUNCTION: COULD BE INVOLVED IN CELL-CELL AND/OR CELL-MATRIX  
 CC INTERACTIONS NECESSARY FOR NORMAL GROWTH CONTROL.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: MAY BE SECRETED BY A NON-CLASSICAL SECRETORY  
 CC PATHWAY.  
 CC -1- TISSUE SPECIFICITY: MAINLY IN STRATIFIED SQUAMOUS EPITHELIUM.  
 CC -1- SIMILARITY: BELONGS TO THE GALAPRIN (S-LECTIN) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L07769; AAA67899.1; -;  
 DR EMBL; U06643; AAA86820.1; -;  
 DR PDB; 1BK2; 04-NOV-98.  
 DR PDB; 2GAL; 04-NOV-98.  
 DR PDB; 3GAL; 04-NOV-98.  
 DR PDB; 4GAL; 04-NOV-98.  
 DR PDB; 5GAL; 04-NOV-98.  
 DR Aarhus/Ghent-2DPAGE; 17; IEF.  
 DR GeneW; HGNC:6568; LGALST.  
 DR MIM; 600615; -;  
 DR InterPro; IPR001079; Galectin.  
 DR Pfam; PF00337; Gal-binding lectin; 1.  
 DR SMART; SM00276; GLECT; 1.  
 DR PROSITE; PS00309; GALAPTIN; 1.  
 KW Galaptin; Lectin; 3D-structure.  
 FT INIT\_MET 0  
 FT BINDING 69 75  
 SQ SEQUENCE 135 AA; 14944 MW; 4E7CEA54036EF806 CRC64;  
 Query Match 44.0%; Score 40; DB 1; Length 135;  
 Best Local Similarity 75.0%; Pred. No. 9.3;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWRHRIPL 9  
 DB 108 HWRHRIPL 115

RESULT 15  
 ID T2SL\_STRAL  
 AC 053608;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Type II restriction enzyme SalI (EC 3.1.21.4) (Endonuclease SalI)  
 DE (R. Salir).  
 GN SALIR.  
 OS Streptomyces albus G.  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
 NX NCBI\_TaxID=1962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95129852; PubMed=7828868;  
 RA Rodicio M.R., Quinton-Jager T., Moran L.S., Slatko B.E., Wilson G.G.;  
 RA "Organization and sequence of the SalI restriction-modification  
 RT system."; Gene 151:167-172(1994).  
 RL Gene 151:167-172(1994).  
 CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE ATTCGAC AND  
 CC CLEAVES AFTER G-1.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give  
 CC specific double-stranded fragments with terminal 5'-phosphates.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U01232; AAA81886.1; -;  
 DR REBASE; 1588; SALI  
 DR Hydrolyase; Endonuclease; Nuclease; Restriction system.  
 KW SEQUENCE 315 AA; 35337 MW; DAYC3CAB8B924F07 CRC64;  
 Query Match 44.0%; Score 40; DB 1; Length 315;  
 Best Local Similarity 57.1%; Pred. No. 22;  
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Search completed: November 27, 2002, 07:18:55  
 Job time : 5.14286 secs

OY 1 HWRHRIPLQIAG 14  
 DB 301 WYWEHRIDLEAAG 314

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 06:27:07 : Search time 8.33977 Seconds  
(without alignments)  
370.599 Million cell updates/sec

Title: US-09-893-615-2

Perfect score: 91

Sequence: 1 WHMRHRIPQLAAGR 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	52.7	62	10	08S754
2	48	52.7	402	10	09FSM9
3	47	51.6	526	5	09UDW3
4	47	51.6	649	17	09V0U1
5	46.5	51.1	354	16	09XA39
6	46.5	51.1	391	9	003969
7	46.5	51.1	802	16	08UW3
8	45.5	50.0	505	10	022336
9	45.5	50.0	505	10	022336
10	45	49.5	516	16	09KYN3
11	45	49.5	516	16	09A9K8
12	45	49.5	595	17	08ZM38
13	45	49.5	649	17	08U137
14	44.5	48.9	4083	3	09C1M7
15	44.5	48.9	270	16	09HXH2
16	44.5	48.9	319	11	0912B6

17	44.5	48.9	343	11	08VCJ6	08vcj6 mus musculus
18	44.5	48.9	356	16	P95276	P95276 mycobacteri
19	44.5	48.9	467	11	099P02	099p02 mus musculus
20	44.5	48.9	483	11	08VDX5	08vdx5 mus musculus
21	44	48.4	522	13	09YGX4	09ygx4 brachydanio
22	44	48.4	636	16	09PDC8	09pdc8 xylella fas
23	43	47.3	152	16	09K0G3	09k0g3 neisseria m
24	43	47.3	152	16	09JYG2	09jyg2 neisseria m
25	43	47.3	186	6	09BGP1	09bgi1 macaca fasc
26	43	47.3	222	16	09ZBF8	09zbf8 streptomyce
27	43	47.3	322	5	08S2E2	08s2e2 drosophila
28	43	47.3	330	16	09A8W9	09a8w9 caulobacter
29	43	47.3	334	2	052445	052445 pseudomonas
30	43	47.3	367	5	09VPE8	09vpe8 drosophila
31	43	47.3	434	16	09KR89	09kr89 vibrio chol
32	43	47.3	444	3	09C197	09c197 amantia mus
33	43	47.3	466	4	09H8V5	09h8v5 homo sapien
34	43	47.3	491	4	09BZB7	09bzb7 homo sapien
35	43	47.3	527	5	09VGT1	09vgt1 drosophila
36	43	47.3	561	11	P70166	P70166 mus musculus
37	43	47.3	566	4	09BZB8	09bzb8 homo sapien
38	43	47.3	716	16	08YL25	08yl25 anabaena sp
39	43	47.3	930	2	09AQSO	09aqso bacillus sp
40	42.5	46.7	639	16	09K4H5	09k4h5 streptomyce
41	42	46.2	207	16	09RW71	09rw71 deinococcus
42	42	46.2	266	2	09RRV1	09rrv1 mycobacteri
43	42	46.2	289	10	042702	042702 chlamydomon
44	42	46.2	292	16	09K3Q1	09k3q1 streptomyce
45	42	46.2	296	2	09RMP1	09rmp1 mycobacteri

## ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	62 AA.
ID 08S754			
AC 08S754:			
DT 01-JUN-2002 (TREMBLrel. 21, Created)			
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DE Hypothetical 7.3 kDa protein.			
GN OSJNB0081F12.14.			
OS Oryza sativa (Rice).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC Ehrhartoideae; Oryzaceae; Oryza.			
OX NCBI_Taxid=4530;			
RC [1]			
RC SEQUENCE FROM N.A.			
RC STRAIN=NIPONBARE;			
RA McCombie W.R., de la Bastide M., Spiegel L., Kirchoff K., Preston R.,			
RA Kuit K., Nascimento L., Bell M., Balija V., Baker J., Vii M.D.,			
RA Zuluveran T., Santos L., Miller B., Cummins D.M., Shah R., King L.,			
RA Bahret A., Yang C., Dike S., O'Shaughnessy A., Palmer L., Dedila N.;			
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone			
RT OSJNB0081F12, from chromosome 10, complete sequence."			
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL: AC090488; AAM01014.1; -			
KW Hypothetical protein.			
SQ SEQUENCE 62 AA; 7306 MW; 4B141DB22AB8DD41 CRC64;			

Query Match	52.7%	Score 48;	DB 10;	Length 62;
Best Local Similarity	53.3%;	Pred. No. 2.5;		
Matches 8;	Conservative 1;	Mismatches 6;	Indels 0;	Gaps 0;

QY 1 WHMRHRIPQLAAGR 15		
DB 45 WHMRHRIPQLAAGR 59		

RESULT 2  
09FSM9

```

ID Q9FSM9 PRELIMINARY: PRT: 402 AA.
AC Q9FSM9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 44.1 kDa protein.
GN H0711G06.21.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Mu J., Zhou B., Chen Z.H., Li Y., Zhu J.J.,
RA Tang Y.S., Zhao Q., Liu Y.L., Huang Y.C., Yu Z., Fan D.L., Chen L.,
RA Weng Q.J., Zhang L., Lu Y.O., Yu S.L., Zhu J., Liu X.H., Hu X.,
RA Lei H.Y., Zhang Y.J., Meng R., Li C., Lu Y., Chen X.C., Zhang Y.,
RA Hu H., Jia P.X., Li T., Qian Y.M., Ying K., Hong G.F.;
RT "Oryza sativa indica (Guangluai14) genomic DNA, chromosome 4, BAC
RT clone: H0711G06 (+H0113C06).";
RN Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL442115; CAC09515.1;
DR InterPro: IPR004324; BT1.
DR InterPro: IPR003880; Ppantne_attach.
DR Pfam: PF03092; BT1; 1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
DR Hypothetical protein.
SQ SEQUENCE 402 AA: 44120 MW: E2A00524DB4CC3EF CRC64;

Query Match 52.7%; Score 48; DB 10; Length 402;
Best Local Similarity 39.1%; Pred. No. 13;
Matches 9; Conservative 1; Mismatches 5; Indels 8; Gaps 1;

OY 1 HMRHRI-----PQLAAG 15
DB 276 WHMHREGHLLSGPIQARPCR 298

RESULT 3
O9U0W3 PRELIMINARY: PRT: 526 AA.
AC O9U0W3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Possible CAMP-dependent protein kinase subunit.
GN I7276.06.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RA Zimmermann W., Wandt R., Ivens A.C., Murphy L., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RN Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-FRIEDLIN;
RA Rajandream M.A., Barrell B.G.;
RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-FRIEDLIN;
RC MEDLINE=96146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RN Genome Res. 8:135-145(1998).
DR EMBL: AL133436; CAB62823.1;
DR InterPro: IPR000595; cNMP_binding.
DR InterPro: IPR001064; Crystallin.
DR SMART: SM00100; cNMP_1.
DR PROSITE: PS00888; cNMP_BINDING_1; UNKNOWN_1.
DR PROSITE: PS00442; cNMP_BINDING_3; 2.
DR PROSITE: PS00223; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
RA Kinase.

```

```

SQ SEQUENCE 526 AA: 57561 MW: 5CA989A7980BDAAB CRC64;
Query Match 51.6%; Score 47; DB 5; Length 526;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 HMRHRIPLQLAAG 14
DB 378 WPMHQIPIMHVG 391

RESULT 4
O9V0U1 PRELIMINARY: PRT: 649 AA.
AC O9V0U1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein PAB1903.
GN PAB1903.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome
RT structure and evolution.";
RN Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ248285; CAB49612.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 649 AA: 75585 MW: C4918F2188E4B1E5 CRC64;

Query Match 51.6%; Score 47; DB 17; Length 649;
Best Local Similarity 61.5%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 HMRHRIPLQLAAG 14
DB 563 NMHGIPKLAVG 575

RESULT 5
O9XA39 PRELIMINARY: PRT: 354 AA.
AC O9XA39;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE putative epoxide hydrolase.
GN SC03574 OR SCH17.08C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RA Seeger K.D., Harris D.;
RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Elchner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;

```

"A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb streptomycetes coelicolor A3(2) chromosome.";  
RT Mol. Microbiol. 21:77-96(1996).

RA HARPER D., BATEMAN A., BROWN S., CHANDRA G., CHEN C.W., COLLINS M.,  
CROMIN A., FRASER A., GODLE A., HIDALGO J., HORNSBY T., HOWARTH S.,  
HUANG C.-H., KIESER T., LAKE L., MURPHY L., OLIVER K., O'NEIL S.,  
RABINOWITZ E., RAJENDRAM M.A., RUTHERFORD K., RUTTER S.,  
SEEVER K., SAUNDERS D., SHARP S., SQUARES R., SQUARES S., TAYLOR K.,  
WARREN T., WIEZORREK A., WOODWARD J., BARRELL B.G., PARKHILL J.,  
HOPWOOD D.A.,  
RT "Complete genome sequence of the model actinomycete Streptomyces  
coelicolor A3(2)."  
RL Nature 417:141-147(2002).  
DR EMBL: AL079353; CAB45554.1; -  
DR HSSP: P34914; 1CR6.  
DR INTERPRO: IPR000073; Abhydrolase.  
DR INTERPRO: IPR000639; Epox\_hydrolase.  
DR INTERPRO: IPR000379; Ser\_estrs\_site.  
DR Pfam: PF00561; abhydrolase; 1.  
DR PRINTS: PR00412; EPOXHYDRLASE.  
KM Hydrolase.  
SQ SEQUENCE 354 AA; 38122 MW; 4F1305498E0FAC13 CRC64;

Query Match 51.1%; Score 46.5; DB 16; Length 354;  
Best Local Similarity 53.3%; Pred. No. 20;  
Matches 8; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 1 WHRRHRIPLQLAG 14  
|:||||:|  
DB 62 WISRHOLPALAAG 76

RESULT 6  
ID 003969 PRELIMINARY; PRT; 391 AA.  
AC 003969;  
DT 01-JUL-1997 (TREMblrel. 04, Created)  
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE INTEGRASE.  
GN INT.  
OS Bacteriophage phi1g1e.  
OC Viruses.  
OC NCBI\_TaxID=52979;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97225795; PubMed=9073065;  
RA KODALIRA K.I., OKI M., KAKIKAWA M., WATANABE N., HIRAKAWA M.,  
YAMADA K., TAKEO A.;  
RT "Genome structure of the lactobacillus temperate phage phi 1g1e: the  
RT whole genome sequence and the putative promoter/repressor system.";  
RL Gene 187:45-53(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA KODALIRA K.;  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL: X98106; CAA66758.1; -  
DR EMBL: X90510; CAA62092.1; -  
DR InterPro: IPR002104; Phage\_integrase.  
DR Pfam: PF00589; Phage\_integrase; 1.  
SQ SEQUENCE 391 AA; 45517 MW; 0763225F809107A21 CRC64;

Query Match 51.1%; Score 46.5; DB 9; Length 391;  
Best Local Similarity 38.1%; Pred. No. 22;  
Matches 8; Conservative 2; Mismatches 4; Indels 7; Gaps 1;  
OY 1 WHRRHRIPLQLAG 14  
|:||||:|

DB 12 WHRRINTIDQVRVINSAG 32

RESULT 7  
ID 080JM3 PRELIMINARY; PRT; 802 AA.  
AC 080JM3;  
DT 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE ABC transporter, membrane spanning protein.  
GN ATGCG ATU5453 OR AGR PAT 665.  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Plasmid AT.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21608550; PubMed=11743193;  
RA WOOD D.W., SETUBAL J.C., KAUL R., MONKS D.E., KITAJIMA J.P.,  
RA OKURA V.K., ZHOU Y., CHEN L., WOOD G.E., ALMEIDA N.F. Jr., WOO L.,  
RA CHEN Y., PAULSEN I.T., EISEN J.A., KARP P.D., BOYEE D. Sr.,  
RA CHAPMAN P., CLENDENNING J., DEATHERAGE G., GILLET N., GRANT C.,  
RA KUTYAVIN T., LEVY R., LI M.-J., MCCLELLAND E., PALMIERI A.,  
RA RAYMOND C., ROUSE G., SAENPHIMMACHAK C., WU Z., ROMERO P., GORDON D.,  
RA ZHANG S., YOO H., TAO Y., BIDDLE P., JUNG M., KRISPAN W., PERRY M.,  
RA GORDON-KAMM B., LIAO L., KIM S., HENDRICK C., ZHOU Z.-Y., DOLAN M.,  
RA CHUMLEY F., TINGEY S.V., TOMB J.-F., GORDON M.P., OLSON M.V.,  
RA NESTER E.W.;  
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
RT C58."  
RL Science 294:2317-2323(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21608551; PubMed=11743194;  
RA GOODNER B., HINKLE G., GATTUNG S., MILLER N., BLANCHARD M.,  
RA GOURLO B., GOLDMAN B.S., CAO Y., ASKENAZI M., HALLING C., MULLIN L.,  
RA HOUNMEL K., GORDON J., VAUDIN M., IARTCHOUK O., EPP A., LIU F.,  
RA WOLLAM C., ALLINGER M., DOUGHTY D., SCOTT C., LAPPAS C., MARKELZ B.,  
RA CIELLO C., CROWELL C., GURSON J., LOMO C., SEAR C., STRUB G.,  
RT "Genome sequence of the plant pathogen and biotechnology agent  
RT Agrobacterium tumefaciens C58."  
RL Science 294:2323-2328(2001).  
DR EMBL: AE008866; AAL46140.1; ALT\_INT.  
DR EMBL: AE007914; AAK90828.1; -  
KM Plasmid; Complete proteome.  
SQ SEQUENCE 802 AA; 86916 MW; 8D4D192A3A2F7D58 CRC64;

Query Match 51.1%; Score 46.5; DB 16; Length 802;  
Best Local Similarity 90.0%; Pred. No. 42;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 2 WHRRHRIPLQ 11  
|:||||:|  
DB 12 WHRRHRIPLQ 20

RESULT 8  
ID 029897 PRELIMINARY; PRT; 211 AA.  
AC 029897;  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE Conserved hypothetical transmembrane protein.  
GN AF0350.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobales;  
OC Archaeoglobaceae; Archaeoglobus.  
OX NCBI\_TaxID=2234;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9398475;  
RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kellavagge A.R., Graham D.E., Kyrildes N.C.,  
RA Fleischman R.D., Douckenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kikness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Relch C.I., McNeill L.K., Badger J.H., Glodek A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,  
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-  
reducing archaeon *Archaeoglobus fulgidus*.";  
RL Nature 390:364-370(1997).  
DR EMBL: AE001080; AAB90884.1; -.  
DR TIGR: AF0350; -.  
DR InterPro: IPR002781; DUF81.  
DR Pfam: PF01925; DUF81; 1.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
SQ SEQUENCE 211 AA; 23303 MW; 927825BC1EFCEB46 CRC64;

Query Match 50.5%; Score 46; DB 17; Length 211;  
Best Local Similarity 59.1%; Pred. No. 15;  
Matches 13; Conservative 1; Mismatches 0; Indels 8; Gaps 3;

Oy 1 WHWRH-RH-----RIPLOL-AAG 14  
Db 25 WHWRLRYNAPCPRIPLQLPAG 46

## RESULT 9

ID 022236 PRELIMINARY; PRT; 505 AA.  
AC 022236;  
DT 01-JAN-1998 (TREMBLrel. 05; Created)  
DT 01-JAN-1998 (TREMBLrel. 05; Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19; Last annotation update)  
DE Hypothetical 58.5 kDa protein.  
GN T32N15.6.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA de la Bastide M.R., Parnell L.D., Kaplan N., Gnoj L., Hameed A.,  
RA Schütz K., Hasegawa A., Gottesman T., Shohdy N., Granat S., Jensen K.,  
RA Johnson A.F., Lochl M., Dedhia N., Martienssen R., McCombie W.R.;  
RT "A. Thaliana BAC T32N15 from Chromosome V."  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC002534; AAB70030.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 505 AA; 58451 MW; 627A823D89032D0 CRC64;

Query Match 50.0%; Score 45.5; DB 10; Length 505;  
Best Local Similarity 57.1%; Pred. No. 40;  
Matches 8; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Oy 1 WHWRH---RIPLOL 11  
Db 22 WHWRLHYRVPGLT 35

## RESULT 10

ID 09KYN3 PRELIMINARY; PRT; 126 AA.  
AC 09KYN3;  
DT 01-OCT-2000 (TREMBLrel. 15; Created)

DT 01-OCT-2000 (TREMBLrel. 15; Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21; Last annotation update)  
DE Putative membrane protein.  
GN SC07352 OR SC9H11.06.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;  
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-A3(2) / M45;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete *Streptomyces*  
coelicolor A3(2).";  
RL Nature 417:141-147(2002).  
DR EMBL: AL356592; CAB92195.1; -.  
SQ SEQUENCE 126 AA; 13263 MW; 2EB8D08378B22D CRC64;

Query Match 49.5%; Score 45; DB 16; Length 126;  
Best Local Similarity 64.7%; Pred. No. 14;  
Matches 11; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Oy 1 WHWRH-RH-----RIPLOL-AAG 15  
Db 88 WDMRRRPAPOQAALGR 104

## RESULT 11

ID 09A9K8 PRELIMINARY; PRT; 516 AA.  
AC 09A9K8;  
DT 01-JUN-2001 (TREMBLrel. 17; Created)  
DT 01-JUN-2001 (TREMBLrel. 17; Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20; Last annotation update)  
DE TrypTOPhan halogenase, putative.  
GN CC0972.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
RA Potocka I., Nelson W.C., Newton A.C., Stephens C., Phake N.D., Ely B.,  
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Utterback T., Tran K., Wolf A., Yamathayan J., Ermolova M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of *Caulobacter crescentus*."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL: AE005774; AAK22956.1; -.  
DR TIGR: CC0972; -.

KW Complete proteome.  
SQ SEQUENCE 516 AA; 57489 MW; 775AE107AD7CF0EC CRC64;

Query Match 49.5%; Score 45; DB 16; Length 516;  
Best Local Similarity 75.0%; Pred. No. 49;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 3 WHRRIPLOLAG 14  
Db 272 WHRRIPLOLRAG 283

Query Match	49.5%	Score 45;	DB 17;	Length 649;
Best Local Similarity	61.5%	Pred. No. 60;		
Matches	8;	Conservative	2;	Mismatches
			3;	Indels
			0;	Gaps
OY	2 HMRHRIPLQLAG	14		
	:        :			

```

RESULT 15
09HXH2
ID 09HXH2 PRELIMINARY; PRT; 270 AA.
AC 09HXH2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Probable transcriptional regulator.
GN PA8830.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
  Pseudomonas.
OX NCBI_TaxID=287;
  [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou
RA Gader R.L., Goltzy R., Tolentino E., Westbrock-Nadman S., Yuan Y.
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig I.T.,
RA Smith K.A., Spencer D.H., Wong C.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:359-364(2000).
CC -1. SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
  REGULATORS.
DR EMBL, AE004800; AAC07217.1; -.
DR InterPro: IPR000005; HTHARAC.
DR Pfam: PF00165; HTH_ARAC; 2.

```

DR PRINTS: PRO0032; HTHARAC.  
DR SMART: SM00342; HTH\_ARAC; 1.  
DR PROSITE: PS00041; HTH\_ARAC\_FAMILY\_1; UNKNOWN\_1.  
DR PROSITE: PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
KW DNA-binding; transcription regulation; Complete proteome.  
SQ SEQUENCE 270 AA; 29965 MW; 9DCA2B27481E0C80 CRC64;

Query Match 48.9%; Score 44.5; DB 16; Length 270;  
Best Local Similarity 50.0%; Pred. No. 33;  
Matches 9; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

QY 2 HWRHRIPL----QLAG 14  
||||:|  
Db 213 HWRHGVRLAALPRLAG 230

Search completed: November 27, 2002, 07:22:47  
Job time : 16.3398 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:17:51 ; Search time 2.72201 Seconds  
(without alignments)  
162.139 Million cell updates/sec

Title: US-09-893-615-2

Sequence: 1 WHRRRRIPLQLAAGR 15

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	15	2	US-08-592-646A-28
2	91	100.0	15	2	US-08-592-646A-48
3	91	100.0	15	4	US-09-165-422-28
4	91	100.0	15	4	US-09-165-422-48
5	91	100.0	37	2	US-08-592-646A-59
6	91	100.0	37	4	US-09-165-422-59
7	86	94.5	15	2	US-08-592-646A-27
8	86	94.5	15	2	US-09-165-422-27
9	80	87.9	15	2	US-08-592-646A-26
10	80	87.9	15	2	US-08-592-646A-29
11	80	87.9	15	4	US-09-165-422-26
12	80	87.9	15	4	US-09-165-422-29
13	76	83.5	15	2	US-08-592-646A-25
14	76	83.5	15	4	US-09-165-422-25
15	72	79.1	15	2	US-08-592-646A-24
16	72	79.1	15	2	US-08-592-646A-30
17	72	79.1	15	4	US-09-165-422-24
18	72	79.1	15	4	US-09-165-422-30
19	69	75.8	15	2	US-08-592-646A-49
20	69	75.8	15	4	US-09-165-422-49
21	68	74.7	15	2	US-08-592-646A-23
22	68	74.7	15	4	US-09-165-422-23
23	63	69.2	15	2	US-08-592-646A-22
24	63	69.2	15	4	US-09-165-422-22
25	61	67.0	15	2	US-08-592-646A-31
26	61	67.0	15	4	US-09-165-422-31
27	59	64.8	15	2	US-08-592-646A-21

#### ALIGNMENTS

28	59	64.8	15	4	US-09-165-422-21	Sequence 21, Appl
29	56	61.5	11	2	US-08-592-646A-7	Sequence 7, Appl
30	56	61.5	11	4	US-09-165-422-7	Sequence 7, Appl
31	56	61.5	15	2	US-08-592-646A-9	Sequence 9, Appl
32	56	61.5	15	4	US-09-165-422-9	Sequence 9, Appl
33	52	57.1	15	2	US-08-592-646A-20	Sequence 20, Appl
34	52	57.1	15	4	US-09-165-422-20	Sequence 20, Appl
35	51	56.0	15	2	US-08-592-646A-43	Sequence 43, Appl
36	51	56.0	15	2	US-08-592-646A-44	Sequence 43, Appl
37	51	56.0	15	4	US-09-165-422-43	Sequence 44, Appl
38	51	56.0	15	4	US-09-165-422-44	Sequence 44, Appl
39	50	54.9	11	2	US-08-592-646A-8	Sequence 8, Appl
40	50	54.9	11	4	US-09-165-422-8	Sequence 8, Appl
41	50	54.9	15	2	US-08-592-646A-10	Sequence 10, Appl
42	50	54.9	15	2	US-08-592-646A-41	Sequence 41, Appl
43	50	54.9	15	2	US-08-592-646A-45	Sequence 45, Appl
44	50	54.9	15	4	US-09-165-422-10	Sequence 10, Appl
45	50	54.9	15	4	US-09-165-422-41	Sequence 41, Appl

RESULT 1  
US-08-592-646A-28  
Sequence 28, Application US/08592646A  
Patent No. 585135

GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MINOTOPIC POLYPEPTIDES OF TOXOPLASMA  
TITLE OF INVENTION: GONDII AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIVE & BRIDGE, PLC  
STREET: P. O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-28

Query Match 100.0%, Score 91; DB 2; Length 15;  
Best Local Similarity 100.0%, Pred. No. 1.8e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHRRRRIPLQLAAGR 15  
DB 1 WHRRRRIPLQLAAGR 15

RESULT 2  
US-08-592-646A-48  
Sequence 48, Application US/08592646A  
Patent No. 5851535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-48

Query Match 100.0%; Score 91; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHMRRIPLOLAQR 15  
Db 1 WHMRRIPLOLAQR 15

RESULT 3  
US-09-165-422-28  
Sequence 28, Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422

FILING DATE: 02-Oct-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-165-422-28

Query Match 100.0%; Score 91; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHMRRIPLOLAQR 15  
Db 1 WHMRRIPLOLAQR 15

RESULT 4  
US-09-165-422-48  
Sequence 48, Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
US-09-165-422-48

Query Match 100.0%; Score 91; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPIQLAGR 15  
DB 1 WHMRHRIPIQLAGR 15

## RESULT 5

US-08-592-646A-59  
Sequence 59, Application US/08592646A

Patent No. 5851535

GENERAL INFORMATION:

APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA

TITLE OF INVENTION: GONDII AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIF & BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,646A

FILING DATE: 26-JAN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 37 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-592-646A-59

Query Match

Best Local Similarity 100.0%; Score 91; DB 2; Length 37;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPIQLAGR 15

DB 10 WHMRHRIPIQLAGR 24

RESULT 6

US-09-165-422-59

Sequence 59, Application US/09165422

Patent No. 6372897

GENERAL INFORMATION:

APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA

TITLE OF INVENTION: GONDII AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIF & BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,646A

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998  
CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 37 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 59:

US-09-165-422-59

Query Match

Best Local Similarity 100.0%; Score 91; DB 4; Length 37;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPIQLAGR 15

DB 10 WHMRHRIPIQLAGR 24

RESULT 7

US-08-592-646A-27

Sequence 27, Application US/08592646A

Patent No. 5851535

GENERAL INFORMATION:

APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA

TITLE OF INVENTION: GONDII AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIF & BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,646A

FILING DATE: 26-JAN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-592-646A-27

Query Match 94.5%; Score 86; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHMRRIPLOLAG 14  
Db 2 WHMRRIPLOLAG 15

RESULT 8

US-09-165-422-27

Sequence 27, Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
CONDIT AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF &amp; BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: 02-Oct-1998

CLASSIFICATION: &lt;unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-165-422-27

Query Match 94.5%; Score 86; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHMRRIPLOLAG 14  
Db 2 WHMRRIPLOLAG 15

RESULT 9

US-08-592-646A-26

Sequence 26, Application US/08592646A  
Patent No. 5851535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
CONDIT AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC

STREET: P.O. Box 19928  
CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: 26-JAN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-592-646A-26

Query Match 87.9%; Score 80; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHMRRIPLOLAG 13  
Db 3 WHMRRIPLOLAG 15

RESULT 10

US-08-592-646A-29

Sequence 29, Application US/08592646A  
Patent No. 5851535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
CONDIT AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF &amp; BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: 26-JAN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-29

Query Match 87.9%; Score 80; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMRHRIPQLAAGR 15  
| | | | | | | | | | | | | | | | |  
DB 1 HMRHRIPQLAAGR 14

RESULT 11  
US-09-165-422-26  
Sequence 26, Application US/09165422  
Patent No. 6372897

GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
CONDIT AND APPLICATIONS

NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998

CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids

TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-165-422-26

Query Match 87.9%; Score 80; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HMRHRIPQLAA 13  
| | | | | | | | | | | | | | | | |  
DB 3 HMRHRIPQLAA 15

RESULT 12  
US-09-165-422-29

Sequence 29, Application US/09165422  
Patent No. 6372897

GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
CONDIT AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 29:

US-09-165-422-29

Query Match 87.9%; Score 80; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMRHRIPQLAAGR 15  
| | | | | | | | | | | | | | | | |  
DB 1 HMRHRIPQLAAGR 14

RESULT 13  
US-08-592-646A-25  
Sequence 25, Application US/08592646A  
Patent No. 5851535

GENERAL INFORMATION:

APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
CONDIT AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-JAN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-592-646A-25

Query Match 83.5%; Score 76; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.3e-06;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHRRRIPQLA 12  
DB 4 WHRRRIPQLA 15

RESULT 14  
US-09-165-422-25  
Sequence 25, Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
GONDII AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIEF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-165-422-25

Query Match 83.5%; Score 76; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.3e-06;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHRRRIPQLA 12  
DB 4 WHRRRIPQLA 15

RESULT 15  
US-08-592-646A-24  
Sequence 24, Application US/08592646A  
Patent No. 3651535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
GONDII AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIEF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-24

Query Match 79.1%; Score 72; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHRRRIPQL 11  
DB 5 WHRRRIPQL 15

Search completed: November 27, 2002, 07:25:10  
Job time : 2.72201 secs

Sequence	Synthetic template
Sequence 15mer2-5/5	
Peptide #6	which b
Peptide #8	which b
Peptide #9	which b
Human cytoskeleton	
Drosophila melanogaster	
Peptide #7	which b
Peptide #4	which b
Peptide #5	which b
Peptide #9	which b
Human ORF2943	
Human hydrolase 11	
Human secreted protein	
Human secreted protein	
Human mature TANG	
Propionibacterium	
Human ORF protein	
Human secreted protein	
Human secreted protein	
Human secreted protein	
Human secreted protein	
Human secreted protein	
Propionibacterium	
Drosophila melanogaster	
C glutamicum protease	
Corynebacterium glutamicum	
Novel secreted protein	
Human gene 31	encodes
Human secreted protein	
Human secreted protein	
Novel human domain	
Human protein AF02	
Staphylococcus aureus	
Human secreted protein	
Bacteriophage 192	

XX Example 2; Page 7; 33pp; French.  
PS  
CC The invention relates to novel peptides which bind to antibodies which  
CC recognise the toxoplasma gondii P30 envelop protein. A peptide library  
CC was generated to express pentadecapeptides on the surface of a  
CC filamentous phage as a fusion protein with the PIII protein. The library  
CC was screened with immobilised anti-T. gondii P30 protein antibody 1E1E7.  
CC Phages which bind this antibody were recovered and amplified by one  
CC round of infection in E. coli. The resultant phages were rescreened with  
CC the immobilised antibody and the second round screen isolated 58  
CC bacterial colonies infected with phage. Of the 58 colonies, phage DNA  
CC from 30 colonies was isolated and sequenced to determine the sequence of  
CC the pentadecapeptide encoded. The peptide sequences AAW12276-86 were  
CC identified. Of the 30 colonies studied, this peptide sequence was  
CC encoded 11 times. A template peptide sequence (AAW12287) corresponding  
CC to the sequence across the phage PIII sequence and putative  
CC pentadecapeptide was used to generate a series of overlapping  
CC pentadecapeptides. These peptides were used to determine the best  
CC peptide sequence which binds the 1E1E7 antibody. Peptides AAW03367-75  
CC were isolated. The new peptides can then be used in diagnostic assays to  
CC detect T. gondii antibodies in a sample or to purify anti-P30 antibodies  
CC or for active immunisation against T. gondii.  
XX  
SQ Sequence 15 AA;  
Query Match 100.0%; Score 15; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WHMRHRIPDLAAGR 15  
DB 1 WHMRHRIPDLAAGR 15  
|||||  
RESULT 2  
AAW71334  
ID AAW71334 standard; peptide; 15 AA.  
XX  
AC AAW71334;  
XX  
DT 25-NOV-1998 (first entry)  
XX  
DE A glycolipid sugar chain peptide.  
XX  
KM Glycolipid sugar chain; inhibit; adhesion; metastasis; cancer cell.  
XX  
OS Synthetic.  
XX  
PN JP10237099-A.  
XX  
PD 08-SEP-1998.  
XX  
PF 26-FEB-1997; 97JP-0042311.  
XX  
PR 26-FEB-1997; 97JP-0042311.  
XX  
PA (IMMO ) IMMUNO JAPAN INC.  
XX  
DR WPI; 1998-537488/46.  
XX  
PT New peptide which reacts specifically with antibody against  
PT glyco-lipid sugar chains - useful for inhibition of cancer  
PT metastasis  
PS Claim 8; Page 3; 7pp; Japanese.  
XX  
XX AAW71332-36 represent glycolipid sugar chain replica peptides. They  
CC react specifically with an antibody against glycolipid sugar chains  
CC and inhibit adhesion and metastasis of cancer cells to a target cell.  
CC The peptides can be used to prevent cancer metastasis.  
XX  
SQ Sequence 15 AA;

Query Match 100.0%; Score 15; DB 19; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WHMRHRIPDLAAGR 15  
DB 1 WHMRHRIPDLAAGR 15  
|||||  
RESULT 3  
AAW94702  
ID AAW94702 standard; peptide; 15 AA.  
XX  
AC AAW94702;  
XX  
DT 22-APR-1999 (first entry)  
XX  
DE Lipoteichoic acid epitope peptide mimic for Mab 96-110.  
XX  
KM Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
KM immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
KM Mab 96-110.  
XX  
OS Staphylococcus sp.  
XX  
PN WO9857994-A2.  
XX  
PD 23-DEC-1998.  
XX  
PF 16-JUN-1998; 98WO-US12402.  
XX  
PR 16-JUN-1997; 97US-0049871.  
XX  
PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
XX  
PI Fischer GW, Schuman RF, Stinson JL, Wong H;  
DR WPI; 1999-095329/08.  
XX  
PT New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT infections caused by gram positive bacteria  
XX  
PS Claim 16; Page 120; 150pp; English.  
XX  
CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic acid  
CC of gram positive bacteria, where the Mab is a chimeric immunoglobulin  
CC comprising at least part of a human immunoglobulin constant region and  
CC at least part of a non-human immunoglobulin variable region having  
CC specificity to lipoteichoic acid of gram positive bacteria. The  
CC antibodies bind to whole bacteria and enhance phagocytosis and killing of  
CC the bacteria and enhance protection from lethal infection. The antibodies  
CC or peptides (encoded by a DNA of the variable region of anti-lipoteichoic  
CC acid antibody or characterised by amino acids corresponding to one or  
CC more of the complementarity determining regions (CDRs) of the variable  
CC region of the antibody) can be used for treating or preventing infections  
CC caused by gram positive bacteria. They can also be used for the diagnosis  
CC of gram positive bacterial infections. The present sequence represents a  
CC specifically claimed lipoteichoic acid epitope peptide mimic that can be  
CC bound by the antibody of the invention. (Mab 96-110).  
XX  
SQ Sequence 15 AA;  
Query Match 100.0%; Score 15; DB 20; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WHMRHRIPDLAAGR 15  
DB 1 WHMRHRIPDLAAGR 15  
|||||



RESULT 4  
 AAB17987 standard; Peptide: 15 AA.  
 XX  
 AC AAB17987;  
 XX  
 DT 31-OCT-2000 (first entry)  
 XX  
 DE Beta-2GPI Ab binding peptide sequence SEQ ID NO:1099.  
 XX  
 KW Modified peptide: therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;  
 KW MMP inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase;  
 KW asthma; thrombosis; pharmaceutical.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200024782-A2.  
 XX  
 PD 04-MAY-2000.  
 XX  
 PE 25-OCT-1999; 99WO-US25044.  
 XX  
 PR 23-OCT-1998; 98US-0105371.  
 PR 22-OCT-1999; 99US-0428082.  
 XX  
 PA (AMGE-) AMGEN INC.  
 PI Feige U, Liu C, Cheetham J, Boone TC;  
 XX  
 DR WPI: 2000-350702/30.  
 XX  
 PT Novel composition of matter comprising an Fc domain and  
 PT pharmacologically active peptides, useful for treating cancer and  
 PT autoimmune diseases -  
 PS  
 PS Claim 39; Page 598; 608pp; English.  
 XX  
 CC The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,  
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4  
 CC where P1, P2, P3, and P4 = are each independently sequences of  
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each  
 CC independently linkers; and a, b, c, d, e, and f = are each independently  
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can  
 CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive  
 CC activities. DNAs, vectors and host cells from the present invention can  
 CC be used for producing pharmaceutical compositions. The compositions are  
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.  
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer  
 CC half-life or incorporate functions such as Fc receptor binding, protein  
 CC A binding, complement fixation, and possibly placental transfer. AAB6443  
 CC to AAB6936 and AAB16955 to AAB18003 represent nucleotide and amino acid  
 CC sequences used in the exemplification of the present invention.  
 CC  
 XX  
 SQ Sequence 15 AA:  
 Query Match 100.0%; Score 15; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-08;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 WHMRHRIPQLAAGR 15  
 |||||||||||||  
 Db 1 WHMRHRIPQLAAGR 15

ID ABB73358 standard; Peptide: 15 AA.  
 XX  
 AC ABB73358;  
 XX  
 DT 05-APR-2002 (first entry)  
 XX  
 DE Exemplary pharmacologically active peptide SEQ ID NO:1097.  
 XX  
 KW Modified peptide: mimetic; Fc domain; fusion; immunoglobulin G; IgG;  
 KW EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 KW  
 XX  
 OS Synthetic.  
 XX  
 PN WO200183525-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PE 02-MAY-2001; 2001WO-US14310.  
 XX  
 PR 03-MAY-2000; 2000US-0563286.  
 XX  
 PA (AMGE-) AMGEN INC.  
 PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;  
 XX  
 DR WPI: 2002-130313/17.  
 XX  
 PT Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases; cancer; rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility -  
 PS  
 PS Claim 39; Page 62; 176pp; English.  
 XX  
 CC The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising  
 CC EPO-mimetic compounds are useful for treating disorders characterised by  
 CC low red blood cell levels such as anaemia. The TPO-mimetic compounds  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention.  
 CC  
 XX  
 SQ Sequence 15 AA:  
 Query Match 100.0%; Score 15; DB 23; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-08;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 WHMRHRIPQLAAGR 15

Db 1 MHWHRIRIPLQLAAGR 15

RESULT 6  
AAW94729  
ID AAW94729 standard; Protein; 19 AA.

XX AAW94729;

DT 22-APR-1999 (first entry)

DE Sequence 15mer 2nd.1 resulting from library panning experiments.

XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;

KW Mab 96-110; panning.

XX Staphylococcus sp.

XX WO9857994-A2.

PD 23-DEC-1998.

PF 16-JUN-1998; 98WO-US12402.

PR 16-JUN-1997; 97US-0049871.

PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PI Fischer GW, Schuman RF, Stinson JL, Wong H;

XX MPI: 1999-095329/08.

DR N-PSDB; AAX05555.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used

PT to develop products for the diagnosis, prevention and treatment of

PT infections caused by gram positive bacteria

XX Example 5; Fig 8; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic

CC acid of gram positive bacteria, where the Mab is a chimeric

CC immunoglobulin comprising at least part of a human immunoglobulin

CC constant region and at least part of a non-human immunoglobulin variable

CC region having specificity to lipoteichoic acid of gram positive bacteria.

CC The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The

CC antibodies or peptides (encoded by a DNA of the variable region of

CC anti-lipoteichoic acid antibody or characterised by amino acids

CC corresponding to one or more of the complementarity determining regions

CC (CDRs) of the variable region of the antibody) can be used for treating

CC or preventing infections caused by gram positive bacteria. They can also

CC be used for the diagnosis of gram positive bacterial infections.

CC Sequences AAW94726-34 represent common peptide sequences resulting from

CC all library panning experiments. Three series of panning experiments were

CC conducted to identify peptide sequences to which antibody of the

CC invention (Mab 96-110) bound strongly. The translated sequences provide

CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX Sequence 19 AA;

XX Query Match 100.0%; Score 15; DB 20; Length 19;

XX Best Local Similarity 100.0%; Pred. No. 1.7e-08;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 MHWHRIRIPLQLAAGR 15

XX DB 3 MHWHRIRIPLQLAAGR 17

XX RESULT 7

XX AAW94710

ID AAW94710 standard; Protein; 19 AA.

XX AAW94710;

DT 22-APR-1999 (first entry)

DE Sequence 15mer 2-8/0 resulting from 15mer library panning experiment.

XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;

KW Mab 96-110; panning.

XX Staphylococcus sp.

XX WO9857994-A2.

PD 23-DEC-1998.

PF 16-JUN-1998; 98WO-US12402.

PR 16-JUN-1997; 97US-0049871.

PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PI Fischer GW, Schuman RF, Stinson JL, Wong H;

XX MPI: 1999-095329/08.

DR N-PSDB; AAX05536.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used

PT to develop products for the diagnosis, prevention and treatment of

PT infections caused by gram positive bacteria

XX Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic

CC acid of gram positive bacteria, where the Mab is a chimeric

CC immunoglobulin comprising at least part of a human immunoglobulin

CC constant region and at least part of a non-human immunoglobulin variable

CC region having specificity to lipoteichoic acid of gram positive bacteria.

CC The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The

CC antibodies or peptides (encoded by a DNA of the variable region of

CC anti-lipoteichoic acid antibody or characterised by amino acids

CC corresponding to one or more of the complementarity determining regions

CC (CDRs) of the variable region of the antibody) can be used for treating

CC or preventing infections caused by gram positive bacterial infections.

CC Sequences AAW94705-22 represent common peptide sequences resulting from a 15mer library

CC panning second experiment. Three series of panning experiments were

CC conducted to identify peptide sequences to which antibody of the

CC invention (Mab 96-110) bound strongly. The translated sequences provide

CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX Sequence 19 AA;

XX Query Match 100.0%; Score 15; DB 20; Length 19;

XX Best Local Similarity 100.0%; Pred. No. 1.7e-08;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 MHWHRIRIPLQLAAGR 15

XX DB 3 MHWHRIRIPLQLAAGR 17

XX RESULT 8

XX AAW94721

XX ID AAW94721 standard; Protein; 19 AA.

XX AC AAW94721;

XX DT 22-APR-1999 (first entry)

XX

```

DE Sequence 15mer2-19/0 resulting from 15mer library panning experiment.
XX
XX Monoclonal antibody: Mab; lipoteichoic acid; gram positive; bacteria;
XX immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX Mab 96-110; panning.
XX
OS Staphylococcus sp.
XX
XX MO9857994-A2.
XX
XX 23-DEC-1998.
XX
XX 16-JUN-1998; 98WO-US12402.
XX
XX 16-JUN-1997; 97US-0049871.
XX
XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
XX WPI: 1999-095329/08.
XX
XX N-PSDB: AAX05547.
XX
XX
XX New antibodies to lipoteichoic acid of gram positive bacteria - used
XX to develop products for the diagnosis, prevention and treatment of
XX infections caused by gram positive bacteria
XX
XX Example 5; Fig 6A-B; 150pp; English.
XX
XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic
XX acid of gram positive bacteria, where the Mab is a chimeric
XX immunoglobulin comprising at least part of a human immunoglobulin
XX constant region and at least part of a non-human immunoglobulin variable
XX region having specificity to lipoteichoic acid of gram positive bacteria.
XX The antibodies bind to whole bacteria and enhance phagocytosis and
XX killing of the bacteria and enhance protection from lethal infection. The
XX antibodies or peptides (encoded by a DNA of the variable region of
XX anti-lipoteichoic acid antibody or characterised by amino acids
XX corresponding to one or more of the complementarity determining regions
XX (CDRs) of the variable region of the antibody) can be used for treating
XX or preventing infections caused by gram positive bacteria. They can also
XX be used for the diagnosis of gram positive bacterial infections.
XX Sequences AAM94705-22 represent sequences resulting from a 15mer library
XX panning second experiment. Three series of panning experiments were
XX conducted to identify peptide sequences to which antibody of the
XX invention (Mab 96-110) bound strongly. The translated sequences provide
XX lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.
XX
XX Sequence 19 AA:
XX
XX Query Match 100.0%; Score 15; DB 20; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 17e-08;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 WHWRRIRIPQLAAGR 15
XX ||||||||||||
XX ||||||||||||
XX DB 3 WHWRRIRIPQLAAGR 17
XX
XX
XX RESULT 9
XX AAM94705
XX ID AAM94705 standard; Protein; 19 AA.
XX
XX AC AAM94705;
XX
XX DT 22-APR-1999 (first entry)
XX
XX DE Sequence 15mer2-1/0 resulting from 15mer library panning experiment.
XX
XX Monoclonal antibody: Mab; lipoteichoic acid; gram positive; bacteria;
XX immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX Mab 96-110; panning.
XX

```

```

OS   Staphylococcus sp.
XX
XX   WO9857994-A2.
XX
XX   23-DEC-1998.
XX
XX   16-JUN-1998; 98WO-US12402.
XX
XX   16-JUN-1998; 97US-0049871.
XX
XX   (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX   Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
XX   WPI; 1999-095329/08.
XX
XX   N-PSDB; AAX05531.
XX
XX   New antibodies to lipoteichoic acid of gram positive bacteria - used
XX   to develop products for the diagnosis, prevention and treatment of
XX   infections caused by gram positive bacteria
XX
XX   Example 5: Fig 6A-B; 150pp; English.
XX
XX   The invention relates to a monoclonal antibody (Mab) to lipoteichoic
XX   acid of gram positive bacteria, where the Mab is a chimeric
XX   immunoglobulin comprising at least part of a human immunoglobulin
XX   constant region and at least part of a non-human immunoglobulin variable
XX   region having specifically to lipoteichoic acid of gram positive bacteria.
XX   The antibodies bind to whole bacteria and enhance phagocytosis and
XX   killing of the bacteria and enhance protection from lethal infection. The
XX   antibodies or peptides (encoded by a DNA of the variable region of
XX   anti-lipoteichoic acid antibody or characterised by amino acids
XX   (CPRS) of the variable region of the complementarity determining regions
XX   or preventing infections caused by gram positive bacteria. They can also
XX   be used for the diagnosis of gram positive bacterial infections.
XX   Sequences AAW94705-22 represent sequences resulting from a 15mer library
XX   panning second experiment. Three series of panning experiments were
XX   conducted to identify peptide sequences to which antibody of the
XX   invention (Mab 96-110) bound strongly. The translated sequences provide
XX   lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.
XX
XX   Sequence 19 AA:
XX
XX   Query Match 100.0%; Score 15; DB 20; Length 19;
XX   Best Local Similarity 100.0%; Prod. No. 1.7e-08;
XX   Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX   QY 1 WHRRRIRPLQLAAGR 15
XX   ||| ||| ||| ||| ||| |||
XX   DB 3 WHRRRIRPLQLAAGR 17
XX
XX   RESULT 10
XX   ID AAW94709
XX   AAW94709 standard; Protein; 19 AA.
XX
XX   AC AAW94709;
XX
XX   DT 22-APR-1999 (first entry)
XX
XX   DE Sequence 15mer2-7/0 resulting from 15mer library panning experiment.
XX
XX   KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;
XX   immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX   Mab 96-110; panning.
XX
XX   OS Staphylococcus sp.
XX
XX   PN WO9857994-A2.
XX
XX   PD 23-DEC-1998.
XX

```

```

PF 16-JUN-1998: 98WO-US12402.
XX
XX 16-JUN-1997: 97US-0049871.
XX
XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
XX WPI; 1999-095329/08.
XX
XX N-PSDB; AAX05535.
XX
PT New antibodies to lipoteichoic acid of gram positive bacteria - used
PT to develop products for the diagnosis, prevention and treatment of
PT infections caused by gram positive bacteria
XX
XX Example 5; Fig 6A-B; 150pp; English.
XX
CC The invention relates to a monoclonal antibody (MAb) to lipoteichoic
CC acid of gram positive bacteria, where the MAb is a chimeric
CC immunoglobulin comprising at least part of a human immunoglobulin
CC constant region and at least part of a non-human immunoglobulin variable
CC region having specificity to lipoteichoic acid of gram positive bacteria.
CC The antibodies bind to whole bacteria and enhance phagocytosis and
CC killing of the bacteria and enhance protection from lethal infection. The
CC antibodies or peptides (encoded by a DNA of the variable region of
CC anti-lipoteichoic acid antibody or characterised by amino acids
CC corresponding to one or more of the complementarity determining regions
CC (CDRs) of the variable region of the antibody) can be used for treating
CC or preventing infections caused by gram positive bacteria. They can also
CC be used for the diagnosis of gram positive bacterial infections.
CC Sequences AAW94703-22 represent sequences resulting from a 15mer library
CC panning second experiment. Three series of panning experiments were
CC conducted to identify peptide sequences to which antibody of the
CC invention (Mab 96-110) bound strongly. The translated sequences provide
CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.
XX
SQ Sequence 19 AA:
Query Match 100.0%; Score 15; DB 20; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHMHRIRIQLAAGR 15
   |||||
DB 3 WHMHRIRIQLAAGR 17

RESULT 11
AAW12287
ID AAW12287 standard; peptide; 37 AA.
XX
XX AAW12287;
AC
XX
XX 15-APR-1997 (first entry)
DT
XX
XX Synthetic template peptide for overlapping peptide series.
DE
XX
XX Toxoplasma gondii; surface protein; antibody; screening; peptide library;
KM diagnostic assay; immunisation; phage; fusion protein; envelop.
XX
XX Synthetic.
OS
XX
XX EP724016-A1.
XX
XX 31-JUL-1996.
XX
XX 29-JAN-1996; 96EP-0420030.
XX
XX 30-JAN-1995; 95FR-0001297.
XX
XX (JOLI/) JOLIVET-REYNAUD C.
XX
XX (INMR) BIO MERIEUX.
XX

```

```

PI Jolivet-Reynaud C;
XX
XX WPI; 1996-343531/35.
XX
XX New polypeptide reactive with anti-P30 antibodies against Toxoplasma
XX gondii useful for diagnosis or immunisation, also new nucleic
XX acid, vectors and transformed cells
XX
XX Example 5; Page 9; 33pp; French.
XX
XX
CC The invention relates to novel peptides which bind to antibodies which
CC recognise the Toxoplasma gondii P30 envelop protein. A peptide library
CC was generated to express pentadecapeptides on the surface of a
CC filamentous phage as a fusion protein with the PIII protein. The library
CC was screened with immobilised anti-T. gondii P30 protein antibody 1E17.
CC Phages which bind this antibody were recovered and amplified by one
CC round of infection in E. coli. The resultant phages were rescreened with
CC the immobilised antibody and the second round screen isolated 58
CC bacterial colonies infected with phage. Of the 58 colonies, phage DNA
CC from 30 colonies was isolated and sequenced to determine the sequence of
CC the pentadecapeptide encoded. The peptide sequences AAW12276-86 were
CC identified. Of the 30 colonies studied, this peptide sequence was
CC encoded once. A template peptide sequence (AAW12287) corresponding to
CC the sequence across the phage PIII sequence and putative
CC pentadecapeptide was used to generate a series of overlapping
CC pentadecapeptides. These peptides were used to determine the best
CC peptide sequence which binds the 1E17 antibody. Peptides AAW03367-75
CC were isolated. The new peptides can then be used in diagnostic assays to
CC detect T. gondii antibodies in a sample or to purify anti-P30 antibodies
CC or for active immunisation against T. gondii.
XX
SQ Sequence 37 AA:
Query Match 100.0%; Score 15; DB 17; Length 37;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHMHRIRIQLAAGR 15
   |||||
DB 10 WHMHRIRIQLAAGR 24

RESULT 12
AAW94708
ID AAW94708 standard; Protein; 19 AA.
XX
XX
XX AAW94708;
AC
XX
XX 22-APR-1999 (first entry)
DT
XX
XX Sequence 15mer2-5/0 resulting from 15mer library panning experiment.
DE
XX
XX Monoclonal antibody; MAb; lipoteichoic acid; gram positive; bacteria;
KM immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
XX Mab 96-110; panning.
XX
XX Staphylococcus sp.
OS
XX
XX WO9857994-A2.
XX
XX 23-DEC-1998.
XX
XX 16-JUN-1998; 98WO-US12402.
XX
XX 16-JUN-1997; 97US-0049871.
XX
XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX Fischer GW, Schuman RF, Stinson JL, Wong H;
XX
XX WPI; 1999-095329/08.
XX
XX N-PSDB; AAX05534.
XX

```

PT New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT Infections caused by gram positive bacteria  
XX  
XX Example 5; Fig 6A-B; 150pp; English.  
XX  
CC The invention relates to a monoclonal antibody (MAb) to lipoteichoic  
CC acid of gram positive bacteria, where the MAb is a chimeric  
CC immunoglobulin comprising at least part of a human immunoglobulin  
CC constant region and at least part of a non-human immunoglobulin variable  
CC region having specificity to lipoteichoic acid of gram positive bacteria.  
CC The antibodies bind to whole bacteria and enhance phagocytosis and  
CC killing of the bacteria and enhance protection from lethal infection. The  
CC antibodies or peptides (encoded by a DNA of the variable region of  
CC anti-lipoteichoic acid antibody or characterised by amino acids  
CC corresponding to one or more of the complementarity determining regions  
CC (CDRs) of the variable region of the antibody) can be used for treating  
CC or preventing infections caused by gram positive bacteria. They can also  
CC be used for the diagnosis of gram positive bacterial infections.  
CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
CC panning second experiment. Three series of panning experiments were  
CC conducted to identify peptide sequences to which antibody of the  
CC invention (Mab 96-110) bound strongly. The translated sequences provide  
CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
CC  
SQ Sequence 19 AA:  
  
Query Match 93.3%; Score 14; DB: 20; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.7e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 2 HMRHRIPQLAAGR 15  
DB 4 HMRHRIPQLAAGR 17  
|||||  
RESULT 13  
AAW03372  
ID AAW03372 standard; peptide: 11 AA.  
XX  
AC AAW03372;  
XX  
DT 15-APR-1997 (first entry)  
XX  
DE Peptide #6 which binds anti-Toxoplasma gondii P30 antibody.  
XX  
KW Toxoplasma gondii; surface protein; antibody; screening; peptide library;  
KW diagnostic assay; immunisation; phage; E.coli.  
XX  
OS Synthetic.  
XX  
PN EP724016-A1.  
XX  
PD 31-JUL-1996.  
XX  
PE 29-JAN-1996; 96EP-0420030.  
XX  
PR 30-JAN-1995; 95FR-0001297.  
XX  
PA (JOLI/) JOLIVET-REYNAUD C.  
PA (INMR) BIO MERIEUX.  
XX  
PI Jolivet-Reynaud C;  
XX  
DR WPI; 1996-343531/35.  
DR N-PSDB; AAT31328.  
XX  
PT New polypeptide reactive with anti-P30 antibodies against Toxoplasma  
PT gondii - useful for diagnosis or immunisation, also new nucleic  
PT acid, vectors and transformed cells  
XX  
PS Claim 7; Page 19; 33pp; French.  
XX

CC Peptides AAW03367-75 are peptides isolated from a peptide library, which  
CC bind to antibody 1E1E7 which recognises the P30 surface protein from  
CC Toxoplasma gondii. The peptide library was constructed by inserting a  
CC nucleic acid sequence encoding a random pentadecapeptide in the P30  
CC envelop protein of a filamentous phage. Antibody 1E1E7 was immobilised  
CC on the bottom of a Petri dish and the expressed peptide library was  
CC overlaid on the antibody. After washing, phages bound on the antibodies  
CC were recovered and the phage amplified by passage through E. coli. The  
CC phages were rescreened by the same method and 58 bacterial colonies were  
CC recovered. Of these, 30 were further studied and the nucleic acid  
CC encoding the pentadecapeptide was isolated. The nucleic acids encoded  
CC the peptides AAW12276-86. A series of overlapping peptides based on the  
CC peptide template AAW12287 was constructed to identify which peptide  
CC sequences bound to the anti-P30 antibody. The peptide scanning isolated  
CC peptides AAW03367-75. The peptides can then be used in diagnostic  
CC assays to detect T. gondii antibodies in a sample or to purify anti-P30  
CC antibodies or for active immunisation against T. gondii.  
CC  
SQ Sequence 11 AA:  
  
Query Match 73.3%; Score 11; DB: 17; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0001;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 5 HRIPQLAAGR 15  
DB 1 HRIPQLAAGR 11  
|||||  
RESULT 14  
AAW03374  
ID AAW03374 standard; peptide: 15 AA.  
XX  
AC AAW03374;  
XX  
DT 15-APR-1997 (first entry)  
XX  
DE Peptide #8 which binds anti-Toxoplasma gondii P30 antibody.  
XX  
KW Toxoplasma gondii; surface protein; antibody; screening; peptide library;  
KW diagnostic assay; immunisation; phage; E.coli.  
XX  
OS Synthetic.  
XX  
PN EP724016-A1.  
XX  
PD 31-JUL-1996.  
XX  
PE 29-JAN-1996; 96EP-0420030.  
XX  
PR 30-JAN-1995; 95FR-0001297.  
XX  
PA (JOLI/) JOLIVET-REYNAUD C.  
PA (INMR) BIO MERIEUX.  
XX  
PI Jolivet-Reynaud C;  
XX  
DR WPI; 1996-343531/35.  
DR N-PSDB; AAT31330.  
XX  
PT New polypeptide reactive with anti-P30 antibodies against Toxoplasma  
PT gondii - useful for diagnosis or immunisation, also new nucleic  
PT acid, vectors and transformed cells  
XX  
PS Claim 7; Page 20; 33pp; French.  
XX  
CC Peptides AAW03367-75 are peptides isolated from a peptide library, which  
CC bind to antibody 1E1E7 which recognises the P30 surface protein from  
CC Toxoplasma gondii. The peptide library was constructed by inserting a  
CC nucleic acid sequence encoding a random pentadecapeptide in the P30  
CC envelop protein of a filamentous phage. Antibody 1E1E7 was immobilised  
CC on the bottom of a Petri dish and the expressed peptide library was  
CC overlaid on the antibody. After washing, phages bound on the antibodies

CC were recovered and the phage amplified by passage through *E. coli*. The  
CC phages were rescreened by the same method and 58 bacterial colonies were  
CC recovered. Of these, 30 were further studied and the nucleic acid  
CC encoding the pentadecapeptide was isolated. The nucleic acids encoded  
CC the peptides AMW12276-86. A series of overlapping peptides based on the  
CC peptide template AMW12287 was constructed to identify which peptide  
CC sequences bound to the anti-P30 antibody. The peptide scanning isolated  
CC peptides AAW03367-75. The peptides can then be used in diagnostic  
CC assays to detect *T. gondii* antibodies in a sample or to purify anti-P30  
CC antibodies for active immunisation against *T. gondii*.  
XX  
SQ Sequence 15 AA;  
Query Match 73.3%; Score 11; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 HRIPQLAAGR 15  
DB 1 HRIPQLAAGR 11  
RESULT 15  
ABB83482  
ID ABB83482 standard; Protein; 529 AA.  
XX  
AC ABB83482;  
XX  
DT 30-SEP-2002 (first entry)  
XX  
DE Human cytoskeleton-associated protein, CSAP-11.  
XX  
KW Human; cytoskeleton-associated protein; CSAP; CSAP-11;  
KW cell proliferative disorder; viral infection; neurological disorder;  
KW transgenic animal; antihypertensive; antiparasitic; antiinflammatory;  
KW virucide; anticonvulsant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; cytoskeletal.  
XX  
OS Homo sapiens.  
XX  
PN W0200253719-A2.  
XX  
PD 11-JUL-2002.  
XX  
PF 04-JAN-2002; 2002WC-US00178.  
XX  
PR 04-JAN-2001; 2001US-260085P.  
PR 13-FEB-2001; 2001US-268554P.  
PR 14-FEB-2001; 2001US-269111P.  
PR 23-FEB-2001; 2001US-271211P.  
XX  
PA (INCYTE GENOMICS INC.  
XX  
PI Lu DAM, Baughn MR, Yao MG, Ding L, Honchell CD, Yue H, Tang YT;  
PI Warren BA, Duggan BM, Xu Y, Walla NK, Griffin JA, Stewart EA;  
PI Gandhi AR, Khan FA, Thangavelu K, Ison CH, Azimzal Y, Hattalia AJA;  
PI Gietzen KJ, Lai PG, Sanjanwala MM, Elliott VS;  
XX  
DR WPI: 2002-583611/62.  
DR N-PSDB; ABB85320.  
XX  
XX  
PT Novel isolated human cytoskeleton-associated protein for diagnosing,  
PT treating or preventing atherosclerosis, psoriasis, leukemia, epilepsy,  
PT ischemic cerebrovascular disease, cerebral neoplasms and Alzheimer's  
PT disease  
XX  
PS Claim 1; Page 134-135; 167pp; English.  
XX  
CC The present sequence is the protein sequence for a human  
CC cytoskeleton-associated protein (CSAP). The CSAP and its coding sequence  
CC are useful in the diagnosis, treatment and prevention of a cell  
CC proliferative disorder such as actinic keratosis, atherosclerosis,  
CC psoriasis, primary thrombocythaemia, leukaemia; a viral infection such as

CC those caused by adenoviruses (acute respiratory disease, pneumonia),  
CC arenaviruses (lymphocytic choriomeningitis); and a neurological disorder  
CC such as epilepsy, ischaemic cerebrovascular disease, stroke, cerebral  
CC neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease or  
CC amyotrophic lateral sclerosis. The CSAP coding sequence is also useful  
CC for creating knock out or knock in humanised animals or transgenic  
CC animals to model human diseases.  
XX  
SQ Sequence 529 AA;  
Query Match 46.7%; Score 7; DB 23; Length 529;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 9 LQLAAGR 15  
DB 332 LQLAAGR 338  
Search completed: November 27, 2002, 07:27:57  
Job time : 7.32819 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:25:20 : Search time 2.25869 Seconds  
(without alignments)  
638.431 Million cell updates/sec

Title: US-09-893-615-2

Perfect score: 15

Sequence: 1 MWMRRRIPLQLAAGR 15

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	40.0	82	2	BB3740
2	6	40.0	111	2	H91187
3	6	40.0	142	2	SS7954
4	6	40.0	203	2	F69381
5	6	40.0	220	2	S60210
6	6	40.0	254	2	H69057
7	6	40.0	309	2	AB3446
8	6	40.0	330	2	D87068
9	6	40.0	334	2	D95982
10	6	40.0	342	2	S50400
11	6	40.0	344	2	D75486
12	6	40.0	347	2	JCS788
13	6	40.0	367	1	DJPS3P
14	6	40.0	413	2	JX0267
15	6	40.0	413	2	A54968
16	6	40.0	440	2	S54981
17	6	40.0	440	2	F70792
18	6	40.0	450	2	F69880
19	6	40.0	492	2	AB3055
20	6	40.0	510	2	C69753
21	6	40.0	522	2	A98231
22	6	40.0	536	2	A34901
23	6	40.0	558	2	T50742
24	6	40.0	585	2	H83199
25	6	40.0	603	2	A87342
26	6	40.0	666	2	T40172
27	6	40.0	691	2	T45933
28	6	40.0	778	2	S41332
29	6	40.0	803	2	T18738

30	6	40.0	1111	2	T01239	hypothetical prote
31	6	40.0	1144	2	A81983	probable DNA-direc
32	6	40.0	1144	2	H81037	DNA polymerase III
33	6	40.0	1937	2	T03224	probable polyketid
34	5	33.3	18	2	PT0239	Ig heavy chain CDR
35	5	33.3	26	2	H90667	hypothetical prote
36	5	33.3	29	2	S03947	hydrogen dehydroge
37	5	33.3	65	2	AF2770	hypothetical prote
38	5	33.3	70	2	A69842	hypothetical prote
39	5	33.3	77	2	A83548	transcription regu
40	5	33.3	79	2	T20201	hypothetical prote
41	5	33.3	81	2	AG0142	molybdopterin (mpt
42	5	33.3	81	2	T10320	hypothetical prote
43	5	33.3	84	2	E70841	hypothetical prote
44	5	33.3	93	2	D66497	hypothetical prote
45	5	33.3	93	2	A72126	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

BB3740  
hypothetical protein BH0722 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: BB3740  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H.  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: BB3740  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-82 <STO>  
A:Cross-References: GB:BA001509; GB:BA000004; MID:g10173176; PIDN:BA804441.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH0722

Query Match 40.0%; Score 6; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAAG 14  
DB 18 LQLAAG 23

##### RESULT 2

H91187  
hypothetical protein ECS4472 [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: H91187  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: H91187  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-111 <HAV>  
A:Cross-References: GB:BA000007; PIDN:BA837895.1; PID:g13363947; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS4472

Query Match 40.0%; Score 6; DB 2; Length 111;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 PLOLAA 13  
|||||  
Db 89 PLOLAA 94

## RESULT 3

S57954

probable molybdenum-pterin-binding-protein - Azotobacter vinelandii

C:Species: Azotobacter vinelandii

C&gt;Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999

C:Accession: S57954

R:Mouncey, N.J.; Mitchenell, L.A.; Pau, R.N.

submitted to the EMBL Data Library, June 1995

A:Description: Mutational analysis of genes of the mod locus involved in molybdate trans

A:Reference number: S57954

A:Accession: S57954

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-142 &lt;MOU&gt;

A:Cross-references: EMBL:249882; NID:g899220; PIDN:CAA90038.1; PID:g899221

Query Match 40.0%; Score 6; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAAG 14  
|||||  
Db 47 LQLAAG 52

## RESULT 4

F69381

flagellin (flaB1-1) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C&gt;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 26-Aug-1999

C:Accession: F69381

R:Klen, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sultson, G.G.; Gill, S.; Kirkness, E.F.

; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kahe, B.P.; Sykes, S.

Smith, H.O.; Moese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MID:98049343; PMID:9389475

A:Accession: F69381

A&gt;Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-203 &lt;KLE&gt;

A:Cross-references: GB:AE001030; GB:AE000782; NID:g2689353; PIDN:AAB90186.1; PID:g264953

C:Superfamily: archaeal flagellin

Query Match 40.0%; Score 6; DB 2; Length 203;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAAG 14  
|||||  
Db 85 LQLAAG 90

## RESULT 5

S60210

foma protein - Streptomyces wedmorensis

C:Species: Streptomyces wedmorensis

C&gt;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 11-Jan-2002

C:Accession: S60210

R:Hidaka, T.; Goda, M.; Kuzuyama, T.; Takel, N.; Hidaka, M.; Seto, H.

Mol. Gen. Genet. 249, 274-280, 1995

A:Title: Cloning and nucleotide sequence of fosfomycin biosynthetic genes of Streptomyces

A:Reference number: S60207; MID:96091152; PMID:7500951

A:Accession: S60210

A&gt;Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-220 <HID>  
A:Cross-references: EMBL:038561; NID:91060999; PID:g1061005  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
C:Genetics:  
A:Gene: foma

Query Match 40.0%; Score 6; DB 2; Length 220;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 PLOLAA 13  
|||||  
Db 53 PLOLAA 58

## RESULT 6

H69057

hypothetical protein MTH1434 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C&gt;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Mar-2000

C:Accession: H69057

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.

; Liu, D.; Spadofora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jilwan,

Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MID:98037514; PMID:9371463

A:Accession: H69057

A&gt;Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-254 &lt;MTH&gt;

A:Cross-references: GB:AE000905; GB:AE000666; NID:g2622541; PIDN:AAB85909.1; PID:g262

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1434

A:Start codon: GTG

C:Superfamily: Methanobacterium thermoautotrophicum hypothetical protein MTH1434

Query Match 40.0%; Score 6; DB 2; Length 254;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAAG 14  
|||||  
Db 196 LQLAAG 201

## RESULT 7

AB3446

hypothetical exported protein BMEI1552 (imported) - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C&gt;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002

C:Accession: AB3446

R:DelVecchio, V.G.; Kapetral, V.; Redkar, R.J.; Patra, G.; Wujer, C.; Los, T.; Ivanov

; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lec

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit

A:Reference number: AD3252; PMID:11756688

A:Accession: AB3446

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-309 &lt;KUR&gt;

A:Cross-references: GB:AE008917; PIDN:AAL52733.1; PID:g17983563; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI1552

A:Map position: I

Query Match 40.0%; Score 6; DB 2; Length 309;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RPLQL 11



Db 251 RIRPL 256

# RESULT 8

hypothetical protein ML1274 [imported] - Mycobacterium leprae  
 C:Species: Mycobacterium leprae  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: D87068  
 R: Cole, S.T.; Eigemeier, K.; Parthill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A: Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
 A: Title: Massive gene decay in the leprosy bacillus.  
 A: Reference number: A86909; MUID:21128732; PMID:11234002  
 A: Accession: D87068  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-330 <STO>  
 A: Cross-references: GB:AL450380; NID:913093207; PIDN:CAC31655.1; GSPDB:GN00147  
 C: Genetics:  
 A: Gene: ML1274

Query Match 40.0%; Score 6; DB 2; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RIRPL 9  
 Db 121 RIRPL 126

# RESULT 9

hypothetical exported protein [imported] - Sinorhizobium meliloti (strain 1021) magaplas  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C:Accession: D95982  
 R: Finn, T.M.; Weldner, S.; Wong, K.; Buhrmaster, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A: Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing endo  
 A: Reference number: A95842; MUID:21396508; PMID:11481431  
 A: Accession: D95982  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-334 <KUR>  
 A: Cross-references: GB:AL591985; PIDN:CAC49524.1; PID:915141011; GSPDB:GN00167  
 A: Experimental source: strain 1021, megaplasmid psymb  
 R: Galibert, F.; Finn, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A: Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hepaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welis, D.H.; Wong, K.; Yeh, K.  
 A: Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A: Reference number: A96039; MUID:21368234; PMID:11474104  
 A: Contents: annotation  
 C: Genetics:  
 A: Gene: SMB20998  
 A: Genome: plasmid

Query Match 40.0%; Score 6; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RIRPL 11  
 Db 276 RIRPL 281

# RESULT 10

S50400  
 hypothetical protein YMR144w - yeast (Saccharomyces cerevisiae)  
 N: Alternate names: hypothetical protein YK375.13  
 C: Species: Saccharomyces cerevisiae  
 C: Date: 13-Jan-1995 #sequence\_revision 24-Feb-1995 #text\_change 19-Apr-2002  
 C: Accession: S50400  
 R: Badcock, K.; Churcher, C.  
 submitted to the EMBL Data Library, December 1994  
 A: Reference number: S50388  
 A: Accession: S50400  
 A: Molecule type: DNA  
 A: Residues: 1-342 <RAD>  
 A: Cross-references: EMBL:247071; NID:9606429; PIDN:CA87358.1; PID:9606442; GSPDB:GNC  
 C: Genetics:  
 A: Gene: MIPS:YMR144w  
 A: Cross-references: SGD:S0004752  
 A: Map position: 13R

Query Match 40.0%; Score 6; DB 2; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQDAG 14  
 Db 274 LQDAG 279

# RESULT 11

probable UDP-glucose 4-epimerase - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: D75486  
 R: White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
 , M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A: Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A: Reference number: A75250; MUID:20036896; PMID:10567266  
 A: Accession: D75486  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-344 <WHI>  
 A: Cross-references: GB:AE001927; GB:AE000513; NID:96458409; PIDN:AAF10287.1; PID:9645  
 A: Experimental source: strain R1  
 C: Genetics:  
 A: Gene: DR0711  
 A: Map position: 1  
 C: Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homolc

Query Match 40.0%; Score 6; DB 2; Length 344;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 QLAAG 15  
 Db 210 QLAAG 215

# RESULT 12

tsec-1 protein (A and B) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 24-Jan-1998 #sequence\_revision 13-Mar-1998 #text\_change 05-Nov-1999  
 C:Accession: J05788; J05789  
 R: Chen, L.; Sato, M.; Inoko, H.; Kimura, M.  
 Biochem. Biophys. Res. Commun. 240, 261-268, 1997  
 A: Title: Molecular cloning and analysis of novel cDNAs specifically expressed in adu  
 A: Reference number: J05788; MUID:98049818; PMID:9388464  
 A: Accession: J05788  
 A: Molecule type: mRNA  
 A: Residues: 1-347 <CHE>  
 A: Cross-references: DDBJ:AB000619; NID:92749770; PIDN:BAA24108.1; PID:d1025015; PID:9

A:Experimental source: testis  
A:Accession: J05789  
A:Molecule type: mRNA  
A:Residues: 1-347 <CH2>  
A:Cross-references: DDBJ:AB000619; NID:g2749770; PIDN:BAA24108.1; PID:dl025015; PID:g274  
A:Experimental source: testis  
C:Comment: This protein plays a role in mammalian spermatogenesis.  
F:38-80,201-332/Domain: coiled-coil #status predicted <CDC>

Query Match 40.0%; Score 6; DB 2; Length 347;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 LQLAG 15  
|||||  
DB 36 LQLAG 41

## RESULT 13

DJPS3P

CNA-directed DNA polymerase (EC 2.7.7.7) III beta chain - Pseudomonas putida  
C:Species: Pseudomonas putida  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 11-Jun-1999  
C:Accession: JY0002  
R:Fujita, M.Q.; Yoshikawa, H.; Ogasawara, N.  
Mol. Gen. Genet. 215, 381-387, 1989  
A:Title: Structure of the dnaA region of Pseudomonas putida: conservation among three ba  
A:Reference number: JY0002; MUID:89218947; PMID:2540413  
A:Accession: JY0002  
A:Molecule type: DNA  
A:Residues: 1-367 <FUJ>  
A:Cross-references: GB:X14791; NID:g45689; PIDN:CAJ2894.1; PID:g45691  
C:Genetics:  
A:Gene: dnaA  
C:Superfamily: DNA-directed DNA polymerase III beta chain  
C:Keywords: DNA replication initiation; nucleotidyltransferase

Query Match 40.0%; Score 6; DB 1; Length 367;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAG 14  
|||||  
DB 284 LQLAG 289

## RESULT 14

JX0267

alpha-1-antiprotease S-1 precursor - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Jun-2000  
C:Accession: JX0267  
R:Salto, A.; Sanojara, H.  
J. Biochem. 113, 456-461, 1993  
A:Title: Rabbit plasma alpha-1-antiprotease S-1: cloning, sequencing, expression, and  
A:Reference number: JX0267; MUID:93293795; PMID:8514734  
A:Accession: JX0267  
A:Molecule type: mRNA  
A:Residues: 1-413 <SAI>  
A:Cross-references: GB:D16104; NID:g286191; PIDN:BAA03678.1; PID:g303762  
A:Experimental source: liver  
A:Note: Part of this sequence, including the amino end of the mature protein, was confir  
C:Superfamily: antithrombin III  
C:Keywords: glycoprotein  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-413/Product: alpha-1-antiprotease S-1 #status experimental <MAT>  
F:65,102,266/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 6; DB 2; Length 413;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAG 14

DB 129 LQLAG 134  
|||||

## RESULT 15

A54968

alpha-1-antitrypsin precursor - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 16-Jul-1999  
C:Accession: A54968  
R:Ray, B.K.; Gao, X.; Ray, A.  
J. Biol. Chem. 269, 22080-22086, 1994  
A:Title: Expression and structural analysis of a novel highly inducible gene encoding  
A:Reference number: A54968; MUID:94350954; PMID:8071331  
A:Accession: A54968  
A:Molecule type: preliminary  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-413 <RAY>  
A:Cross-references: GB:L12139; NID:g405551; PIDN:AA57133.1; PID:g601905  
C:Superfamily: antithrombin III

Query Match 40.0%; Score 6; DB 2; Length 413;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAG 14  
|||||  
DB 129 LQLAG 134

Search completed: November 27, 2002, 07:31:20  
Job time : 6.25869 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:36:38 ; Search time 1259.04 Seconds  
(without alignments)  
7350.598 Million cell updates/sec

Title: US-09-893-615-88

Sequence: 1 CAAATGTTCTCTCCACTC.....GCACATGCTGGAATAAGA 318

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_com:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_com:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_un:\*  
28: em\_vl:\*  
29: em\_hg\_hum:\*  
30: em\_hg\_inv:\*  
31: em\_hg\_mus:\*  
32: em\_hg\_other:\*  
33: em\_hg\_mus:\*  
34: em\_hg\_pln:\*  
35: em\_hg\_rtd:\*  
36: em\_hg\_mam:\*  
37: em\_hg\_vrt:\*  
38: em\_sy:\*  
39: em\_hggo\_hum:\*  
40: em\_hggo\_mus:\*  
41: em\_hggo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	303.6	95.5	321	10	MMIG013
2	301.6	94.8	318	10	AF178620
3	301	94.7	318	10	AF178619
4	300.4	94.5	321	6	AR096128
5	300.4	94.5	321	6	AR210527
6	300.4	94.5	384	6	AR000010
7	300.4	94.5	384	6	AR060923
8	300.4	94.5	384	6	AR211055
9	300.4	94.5	384	6	AX032417
10	300.4	94.5	384	6	BD004717
11	300.4	94.5	9209	6	AR000007
12	300.4	94.5	9209	6	AR060920
13	300.4	94.5	9209	6	AR211052
14	300.4	94.5	9209	6	AX032414
15	300.4	94.5	9209	6	BD004714
16	300.4	94.5	18986	6	AR051652
17	300.4	94.5	18986	6	AR092290
18	298.8	94.0	384	6	AR015962
19	298.8	94.0	1370	10	MUSICKAAR
20	297.2	93.5	9209	6	AR015961
21	295.6	93.0	318	6	AR096196
22	295.6	93.0	318	6	AR210595
23	295.6	93.0	335	6	AR096195
24	295.6	93.0	335	6	AR210594
25	295.6	93.0	403	6	IO9200
26	295.6	93.0	403	10	MUSICKCNA
27	295.6	93.0	477	10	S76823
28	294.2	92.5	345	10	MUSICKVA2A
29	294	92.5	403	6	IO5923
30	292.6	92.0	309	10	MMIGVKPA
31	292.4	91.9	302	10	MUSICKAPE
32	290.8	91.4	350	10	AY058906
33	290.8	91.4	384	10	MMU05285
34	290.8	91.4	403	6	IO8812
35	290.8	91.4	732	6	AS1863
36	290.8	91.4	732	6	AS3237
37	290.8	91.4	732	6	AR085831
38	290.8	91.4	732	6	AR182961
39	289.4	91.0	318	12	AF277092
40	289.4	91.0	765	6	AX057984
41	289.4	91.0	1239	6	AX057945
42	289.4	91.0	1280	6	AX057947
43	288.2	90.6	435	6	A17967
44	288	90.6	360	10	AF029237
45	287.6	90.4	705	6	AB3197

## ALIGNMENTS

RESULT 1  
MMIG013  
LOCUS  
DEFINITION MMIG013 321 bp mRNA linear ROD 06-AUG-1995  
region: Mouse hybridoma 2d3 mRNA for Immunoglobulin kappa light chain V  
ACCESSION X58586 Y00794  
VERSION X58586.1 GI:51562  
KEYWORDS Ig kappa light chain; Ig variable region; immunoglobulin.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 321)  
AUTHORS Meek, K.  
TITLE Direct Submission

JOURNAL Submitted (09-JAN-1989) K. Meek, University of Texas Southwestern Medical School, 5323 Harry Hines Blvd. Microbiology, Dallas Texas 75235, USA

REFERENCE 2 (bases 1 to 321)

AUTHORS Meek, K., Hasemann, C., Pollok, B., Alkan, S.S., Bralt, M., Slaoui, M., Urdain, J., and Capra, J.D.

TITLE Structural characterization of antidiabetic antibodies. Evidence that Ab2s are derived from the germ-line differently than Ab1s

JOURNAL J. Exp. Med. 169 (2), 519-533 (1989)

MEDLINE 89094248

PUBMED 2492056

FEATURES

Source Location/Qualifiers

1..321

/organism="Mus musculus"

/strain="BALB/c"

/db\_xref="taxon:10090"

/cell\_line="Hybridoma 2D3-K"

1..321

/gene="IG kappa light chain"

<1..>321

/gene="IG kappa light chain"

/note="variable region"

/evidence="experimental"

<1..>321

/gene="IG kappa light chain"

/note="variable region"

/product="IG kappa light chain"

/note="variable region"

CDS

<1..>321

/gene="IG kappa light chain"

/product="IG kappa light chain"

/note="variable region"

/db\_xref="GI:938245"

/protein\_id="CA41461.1"

/translation="QIVLSQSPAILASPEKYVMTCRASSSVYMHVYQKRGSSPK PWISATSNLASGVPARFSGSGTSYSLTISRYEADNATYYCHOWSSNPPTGGGTR LEIKR"

V\_region

<1..>321

/gene="IG kappa light chain"

/product="IG kappa light chain"

/note="variable region"

BASE COUNT 78 a 92 c 80 g 71 t

ORIGIN

Query Match 95.5% Score 303.6; DB 10; Length 321; Best Local Similarity 97.2%; Pred. NO. 3.8e-90; Matches 309; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CAATTGTTCTCTCCAGTCTCCAGCAATCCTGTCGATCCAGGGGAAAGTCA 60

DB 1 CAATTGTTCTCTCCAGTCTCCAGCAATCCTGTCGATCCAGGGGAAAGTCA 60

QY 61 ATGACTTGACAGGGCAGCTCAAGTGAATTAATGACATGCTGTACACAGCAAGCCAGA 120

DB 61 ATGACTTGACAGGGCAGCTCAAGTGAATTAATGACATGCTGTACACAGCAAGCCAGA 120

QY 121 TCTCTCCCAAAACCTGGATTTCGCCACATCCAACTGGCTTCGAGTCCCTGCTGC 180

DB 121 TCTCTCCCAAAACCTGGATTTCGCCACATCCAACTGGCTTCGAGTCCCTGCTGC 180

QY 181 TTGAGTGGACATGGGTGGAGCTCTTACTCTCTCAACAATCAGACAGAGGAGCTGAA 240

DB 181 TTGAGTGGACATGGGTGGAGCTCTTACTCTCTCAACAATCAGAGTGGAGGCTGAA 240

QY 241 GATGCTGCACATTAATTAATGCTGACAGCAGTGAAGTAAACCCCAACGTTGGAGGGGG 300

DB 241 GATGCTGCACATTAATTAATGCTGACAGCAGTGAAGTAAACCCCAACGTTGGAGGGGG 300

QY 301 ACCATGCTGCAATTAAGA 318

DB 301 ACCAAGCTGGAATTAAGA 318

RESULT 2

AF178620 318 bp mRNA linear ROD 22-MAY-2000

LOCUS AF178620

DEFINITION Mus musculus 39-9 Immunoglobulin light chain variable region mRNA,

ACCESSION partial cds.

VERSION AF178620

KEYWORDS GI:5853223

SOURCE

ORGANISM Mus musculus.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 318)

AUTHORS Putterman, C., Deocharan, B. and Diamond, B.

TITLE Molecular analysis of the autoantibody response in peptide-induced autoimmunity

JOURNAL J. Immunol. 164 (5), 2542-2549 (2000)

MEDLINE 20143847

PUBMED 10679092

REFERENCE 2 (bases 1 to 318)

AUTHORS Putterman, C., Deocharan, B. and Diamond, B.

TITLE Direct Submission

JOURNAL Submitted (17-AUG-1999) Rheumatology/Medicine, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA

FEATURES

Source Location/Qualifiers

1..318

/organism="Mus musculus"

/strain="BALB/c"

/db\_xref="taxon:10090"

/cell\_line="39-9"

/cell\_type="hybridoma"

/note="from peptide-immunized mouse"

<1..>318

/note="IG kappa: anti-peptide antibody"

/product="Immunoglobulin light chain variable region"

/codon\_start=1

/protein\_id="A554373.1"

/db\_xref="GI:585324"

/translation="QIVLRSQSPAILASPEKYVMTCRASSSVYMHVYQKRGSSPK PWIATSNLASGVPARFSGSGTSYSLTISRYEADNATYYCHOWSSNPPTGGGTR LEIKR"

CDS

<1..>318

/gene="IG kappa: anti-peptide antibody"

/product="Immunoglobulin light chain variable region"

/db\_xref="GI:585324"

/protein\_id="A554373.1"

/translation="QIVLRSQSPAILASPEKYVMTCRASSSVYMHVYQKRGSSPK PWIATSNLASGVPARFSGSGTSYSLTISRYEADNATYYCHOWSSNPPTGGGTR LEIKR"

BASE COUNT 75 a 89 c 81 g 73 t

ORIGIN

Query Match 94.8% Score 301.6; DB 10; Length 318; Best Local Similarity 97.2%; Pred. NO. 1.8e-89; Matches 307; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 AATGTTCTCTCCAGTCTCCAGCAATCCTGTCGATCCAGGGGAAAGTCA 62

DB 3 AATGTTCTCTCCAGTCTCCAGCAATCCTGTCGATCCAGGGGAAAGTCA 62

QY 63 GACTTGACAGGGCAGCTCAAGTGAATTAATGACATGCTGTACACAGCAAGCCAGA 122

DB 63 GACTTGACAGGGCAGCTCAAGTGAATTAATGACATGCTGTACACAGCAAGCCAGA 122

QY 123 CTCCCAAAACCTGGATTTCGCCACATCCAACTGGCTTCGAGTCCCTGCTGC 182

DB 123 CTCCCAAAACCTGGATTTCGCCACATCCAACTGGCTTCGAGTCCCTGCTGC 182

QY 183 CAGTGGACATGGGTGGAGCTCTTACTCTCTCAACAATCAGACAGAGGAGCTGAA 242

DB 183 CAGTGGACATGGGTGGAGCTCTTACTCTCTCAACAATCAGAGTGGAGGCTGAA 242

QY 243 TGCTGCCACTTAATTAATGCTGACAGCAGTGAAGTAAACCCCAACGTTGGAGGGGG 302

DB 243 TGCTGCCACTTAATTAATGCTGACAGCAGTGAAGTAAACCCCAACGTTGGAGGGGG 302

QY 303 CATGCTGCAATTAAGA 318

DB 303 CAAGCTGGAATTAAGA 318

RESULT 3

AF178619 318 bp mRNA linear ROD 22-MAY-2000

LOCUS AF178619

DEFINITION Mus musculus 8-3 Immunoglobulin light chain variable region mRNA,

partial cds.  
 ACCESSION AF178619  
 VERSION AF178619.1 GI:5853221  
 KEYWORDS  
 SOURCE  
 ORGANISM Mus musculus.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 318)  
 AUTHORS Puterman, C., Desocharan, B. and Diamond, B.  
 TITLE Molecular analysis of the autoantibody response in peptide-induced autoimmunity  
 JOURNAL J. Immunol. 164 (5), 2542-2549 (2000)  
 MEDLINE 20143847  
 PUBMED 10679092  
 REFERENCE 2 (bases 1 to 318)  
 AUTHORS Puterman, C., Desocharan, B. and Diamond, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-AUG-1999) Rheumatology/Medicine, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA  
 FEATURES  
 source  
 1..318  
 /organism="Mus musculus"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /cell\_line="8-3"  
 /cell\_type="hybridoma"  
 /note="from peptide-immunized mouse"  
 <1..>318  
 /note="1G kappa; anti-peptide antibody"  
 /codon\_start=1  
 /product="immunoglobulin light chain variable region"  
 /protein\_id="AAB54372.1"  
 /db\_xref="GI:5853222"  
 /translation="KIVISQPAILASPGQKVTMTGRASSSVYMLYQKPPSPK PMISATSNLASGVAPRFSGSGSTSYSLTISRVAEDAAATYCCQWSNPFTFGGTK LEIK"  
 CDS  
 BASE COUNT 76 a 92 c 78 g 72 t  
 ORIGIN  
 Query Match 94.7%; Score 301; DB 10; Length 318;  
 Best Local Similarity 96.8%; Pred. No. 2.8e-89;  
 Matches 307; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Oy 2 AAATGTTCTCTCCAGTCTCCAGCAATCCTGTCGATCTCCAGGGAAGATCACA 61  
 Db 2 AAATGTTATCTCCAGTCTCCAGCAATCCTGTCGATCTCCAGGGAAGATCACA 61  
 Oy 62 TGACTTGAGGCGCCAGTCAAGTGAATTAATGATGATGATGATGATGATGATGAT 121  
 Db 62 TGACTTGAGGCGCCAGTCAAGTGAATTAATGATGATGATGATGATGATGATGAT 121  
 Oy 122 CCTCCCCAAACCTGGAATTCCTGCACATCCACCTGCTTCTGAGTCCCTGCTGCT 181  
 Db 122 CCTCCCCAAACCTGGAATTCCTGCACATCCACCTGCTTCTGAGTCCCTGCTGCT 181  
 Oy 182 TCAGTGGAGTGGGCTGCTGGGACCTTACTCTCCACAATGAGAGATGAGGCTGAG 241  
 Db 182 TCAGTGGAGTGGGCTGCTGGGACCTTACTCTCCACAATGAGAGATGAGGCTGAG 241  
 Oy 242 ATGTGTCACCTTATGTCAGCAGTGAAGTGAATACCCACCTGTCGAGAGGAGGA 301  
 Db 242 ATGTGTCACCTTATGTCAGCAGTGAAGTGAATACCCACCTGTCGAGAGGAGGA 301  
 Oy 302 CCATGCTGGAATAGA 318  
 Db 302 CCAGCTGGAATAGA 318  
 RESULT 4  
 AR096128  
 LOCUS AR096128 321 bp DNA linear PAT 08-SEP-2000  
 DEFINITION Sequence 6 from patent US 6005091.

AR096128  
 VERSION AR096128.1 GI:10024649  
 KEYWORDS  
 SOURCE  
 ORGANISM Unknown.  
 Unknown.  
 REFERENCE 1 (bases 1 to 321)  
 AUTHORS Gross, M. Stuart., Nichols, A. John., Padlan, E. Agustin., Patel, A. Haridhai., and Sylvester, D. Robert.  
 TITLE Nucleic acids encoding immunoglobulin domains  
 JOURNAL Patent: US 6005091-A 6 21-DEC-1999;  
 FEATURES  
 source  
 1..321  
 /organism="unknown"  
 BASE COUNT 78 a 92 c 79 g 72 t  
 ORIGIN  
 Query Match 94.5%; Score 300.4; DB 6; Length 321;  
 Best Local Similarity 96.5%; Pred. No. 4.5e-89;  
 Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 Oy 1 CAAATGTTCTCTCCAGTCTCCAGCAATCCTGTCGATCTCCAGGGAAGATCACA 60  
 Db 1 CAAATGTTCTCTCCAGTCTCCAGCAATCCTGTCGATCTCCAGGGAAGATCACA 60  
 Oy 61 ATGACTTGAGGCGCCAGTCAAGTGAATTAATGATGATGATGATGATGATGATGAT 120  
 Db 61 ATGACTTGAGGCGCCAGTCAAGTGAATTAATGATGATGATGATGATGATGATGAT 120  
 Oy 121 TCCGCCCAAAACCTGGAATTCCTGCACATCCACCTGCTTCTGAGTCCCTGCTGC 180  
 Db 121 TCCGCCCAAAACCTGGAATTCCTGCACATCCACCTGCTTCTGAGTCCCTGCTGC 180  
 Oy 181 TTCAGTGGCAGTGGGTCTGGACCTCTTACTCTTCACATCAGCAGAGTGAAGCTGAA 240  
 Db 181 TTCAGTGGCAGTGGGTCTGGACCTCTTACTCTTCACATCAGCAGAGTGAAGCTGAA 240  
 Oy 241 GATGCTGCACCTTATGTCAGCAGTGAAGTGAATACCCACCTGTCGAGAGGAG 300  
 Db 241 GATGCTGCACCTTATGTCAGCAGTGAAGTGAATACCCACCTGTCGAGAGGAG 300  
 Oy 301 ACCATGCTGGAATAGA 318  
 Db 301 ACCATGCTGGAATAGA 318  
 RESULT 5  
 AR210527  
 LOCUS AR210527 321 bp DNA linear PAT 20-JUN-2002  
 DEFINITION Sequence 6 from patent US 6391299.  
 ACCESSION AR210527  
 VERSION AR210527.1 GI:21513278  
 KEYWORDS  
 SOURCE  
 ORGANISM Unknown.  
 Unknown.  
 REFERENCE 1 (bases 1 to 321)  
 AUTHORS Blackburn, M. Neal., Church, W. Robert., Feuerstein, G. Zeev., Gross, M. Stuart., Nichols, A. John., Padlan, E. Agustin., Patel, A. Haridhai., and Sylvester, D. Robert.  
 TITLE Anti-factor IX/Xa antibodies  
 JOURNAL Patent: US 6391299-A 6 21-MAY-2002;  
 FEATURES  
 source  
 1..321  
 /organism="unknown"  
 BASE COUNT 78 a 92 c 79 g 72 t  
 ORIGIN  
 Query Match 94.5%; Score 300.4; DB 6; Length 321;  
 Best Local Similarity 96.5%; Pred. No. 4.5e-89;  
 Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATTGTTCTCTCCAGCTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCTACA 60  
 DB 1 CAAATTGTTCTCTCCAGCTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCTACA 60  
 QY 61 ATGACTTTCAGGGGCGAGCTCAAGTAAATTACATGCAGCTGACAGAGAGCCAGGA 120  
 DB 61 ATGACTTTCAGGGGCGAGCTCAAGTAAATTACATGCAGCTGACAGAGAGCCAGGA 120  
 QY 121 TCTCTCCCGCAAAACCTGGATTTCTGCCACATCCCAACCTGGCTTCTGGAGTCCCTGCTCGC 180  
 DB 121 TCTCTCCCGCAAAACCTGGATTTATGCCACATCCCAACCTGGCTTCTGGAGTCCCTGCTCGC 180  
 QY 181 TTCAGTGGCAGAGTGGTCTGGGACCTCTTACTCTCTCAATCAATCAGCAGAGTGGAGCTGAA 240  
 DB 181 TTCAGTGGCAGAGTGGTCTGGGACCTCTTACTCTCTCAATCAATCAGCAGAGTGGAGCTGAA 240  
 QY 241 GATGCTGCCACTTATTAATCTCCAGAGTGGAGTAAATCCACCCAGCTTCGGAGGGGG 300  
 DB 241 GATGCTGCCACTTATTAATCTCCAGAGTGGAGTAAATCCACCCAGCTTCGGAGGGGG 300  
 QY 301 ACCATGCTGGAATTAAGA 318  
 DB 301 ACCATGCTGGAATTAAGA 318  
 RESULT 6  
 ARO00010 384 bp DNA linear PAT 04-DEC-1998  
 LOCUS Sequence 6 from patent US 5736137.  
 ACCESSION ARO00010  
 VERSION ARO00010.1 GI:3962541  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 384)  
 AUTHORS Anderson,D.R., Hanna,N., Leonard,J.E., Newman,R.A., Reff,M.E. and Rastetter,W.H.  
 TITLE Therapeutic application of chimeric and radiolabeled antibodies to human B lymphocyte restricted differentiation antigen for treatment of B cell lymphoma  
 JOURNAL Patent: US 5736137-A 6 07-APR-1998;  
 FEATURES  
 source 1..384  
 BASE COUNT 92 a 105 c 93 g 94 t  
 ORIGIN  
 Query Match 94.5%; Score 300.4; DB 6; Length 384;  
 Best Local Similarity 96.5%; Pred. No. 4.5e-89;  
 Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 CAAATTGTTCTCTCCAGCTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCTACA 60  
 DB 67 CAAATTGTTCTCTCCAGCTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCTACA 126  
 QY 61 ATGACTTTCAGGGGCGAGCTCAAGTAAATTACATGCAGCTGACAGAGAGCCAGGA 120  
 DB 127 ATGACTTTCAGGGGCGAGCTCAAGTAAATTACATGCAGCTGACAGAGAGCCAGGA 186  
 QY 121 TCTCTCCCGCAAAACCTGGATTTCTGCCACATCCCAACCTGGCTTCTGGAGTCCCTGCTCGC 180  
 DB 121 TCTCTCCCGCAAAACCTGGATTTATGCCACATCCCAACCTGGCTTCTGGAGTCCCTGCTCGC 180  
 QY 181 TTCAGTGGCAGAGTGGTCTGGGACCTCTTACTCTCTCAATCAATCAGCAGAGTGGAGCTGAA 240  
 DB 181 TTCAGTGGCAGAGTGGTCTGGGACCTCTTACTCTCTCAATCAATCAGCAGAGTGGAGCTGAA 240  
 QY 241 GATGCTGCCACTTATTAATCTCCAGAGTGGAGTAAATCCACCCAGCTTCGGAGGGGG 300  
 DB 241 GATGCTGCCACTTATTAATCTCCAGAGTGGAGTAAATCCACCCAGCTTCGGAGGGGG 300  
 QY 301 ACCATGCTGGAATTAAGA 318  
 DB 301 ACCATGCTGGAATTAAGA 318

DB 367 ACCAAGCTGGAATTAAGA 384  
 RESULT 7  
 ARO60923 384 bp DNA linear PAT 29-SEP-1999  
 LOCUS Sequence 6 from patent US 5843439.  
 ACCESSION ARO60923  
 VERSION ARO60923.1 GI:5988614  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 384)  
 AUTHORS Anderson,D.R., Hanna,N., Leonard,J.E., Newman,R.A., Reff,M.E. and Rastetter,W.H.  
 TITLE Therapeutic application of chimeric and radiolabeled antibodies to human B lymphocyte restricted differentiation antigen for treatment of B cell lymphoma  
 JOURNAL Patent: US 5843439-A 6 01-DEC-1998;  
 FEATURES  
 source 1..384  
 BASE COUNT 92 a 105 c 93 g 94 t  
 ORIGIN  
 Query Match 94.5%; Score 300.4; DB 6; Length 384;  
 Best Local Similarity 96.5%; Pred. No. 4.5e-89;  
 Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 CAAATTGTTCTCTCCAGCTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCTACA 60  
 DB 67 CAAATTGTTCTCTCCAGCTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCTACA 126  
 QY 61 ATGACTTTCAGGGGCGAGCTCAAGTAAATTACATGCAGCTGACAGAGAGCCAGGA 120  
 DB 127 ATGACTTTCAGGGGCGAGCTCAAGTAAATTACATGCAGCTGACAGAGAGCCAGGA 186  
 QY 121 TCTCTCCCGCAAAACCTGGATTTCTGCCACATCCCAACCTGGCTTCTGGAGTCCCTGCTCGC 180  
 DB 121 TCTCTCCCGCAAAACCTGGATTTATGCCACATCCCAACCTGGCTTCTGGAGTCCCTGCTCGC 180  
 QY 181 TTCAGTGGCAGAGTGGTCTGGGACCTCTTACTCTCTCAATCAATCAGCAGAGTGGAGCTGAA 240  
 DB 181 TTCAGTGGCAGAGTGGTCTGGGACCTCTTACTCTCTCAATCAATCAGCAGAGTGGAGCTGAA 240  
 QY 241 GATGCTGCCACTTATTAATCTCCAGAGTGGAGTAAATCCACCCAGCTTCGGAGGGGG 300  
 DB 241 GATGCTGCCACTTATTAATCTCCAGAGTGGAGTAAATCCACCCAGCTTCGGAGGGGG 300  
 QY 301 ACCATGCTGGAATTAAGA 318  
 DB 367 ACCAAGCTGGAATTAAGA 384  
 RESULT 8  
 AR211055 384 bp DNA linear PAT 20-JUN-2002  
 LOCUS Sequence 6 from patent US 6399061.  
 ACCESSION AR211055  
 VERSION AR211055.1 GI:21514275  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 384)  
 AUTHORS Anderson,D.R., Hanna,N., Newman,R.A., Reff,M.E. and Rastetter,W.H.  
 TITLE Chimeric and radiolabelled antibodies specific to human CD20 antigen and use thereof for treatment of B-cell lymphoma  
 JOURNAL Patent: US 6399061-A 6 04-JUN-2002;  
 FEATURES  
 source 1..384

BASE COUNT 92 a 105 c 93 g 94 t  
ORIGIN /organism="unknown"

Query Match 94.5%; Score 300.4; DB 6; Length 384;  
Best Local Similarity 96.5%; Pred. No. 4.5e-89;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 CAAATGTTCTCTCCAGCTCCAGCAATCTGTGCTCCAGGGGAAAAGGTGACA 60  
DB CAAATGTTCTCTCCAGCTCCAGCAATCTGTGCTCCAGGGGAAAAGGTGACA 126  
61 ATGACTTGAGGGCCAGCTCAAGTGAATTAATCATGCACTGTACAGCAGAGCCAGGA 120  
DB ATGACTTGAGGGCCAGCTCAAGTGAATTAATCATGCACTGTACAGCAGAGCCAGGA 186  
121 TCCTCCCCAAACCTGGATTCTGCACATCAACCTGGCTTGGAGTCCCTGCTGC 180  
DB TCCTCCCCAAACCTGGATTCTGCACATCAACCTGGCTTGGAGTCCCTGCTGC 246

241 GATGCTGCCACTTATTACTGCGAGCTAGTAACCCAGCAGCTTGGAGGGGG 300  
DB GATGCTGCCACTTATTACTGCGAGCTAGTAACCCAGCAGCTTGGAGGGGG 366

301 ACCATGCTGGAATAAGA 318  
DB ACCAAGCTGGAATCAAA 384

RESULT 9  
AX032417  
LOCUS AX032417 384 bp DNA linear PAT 20-SEP-2000  
DEFINITION Sequence 6 from Patent EP1005870.  
ACCESSION AX032417  
VERSION AX032417.1 GI:10279390  
KEYWORDS  
SOURCE unclassified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Rastetter,W.H., Hanna,N., Leonard,J.E., Newman,R.A., Reff,M.E. and Anderson,D.R.  
TITLE Therapeutic application of chimeric antibodies to human b lymphocyte restricted differentiation antigen for treatment of b cell lymphoma  
JOURNAL Patent: EP 1005870-A 6 07-JUN-2000;  
IDEC PHARMA CORP (US)  
FEATURES  
source Location/Qualifiers  
1..384  
/organism="unclassified"  
/db\_xref="taxon:32644"

BASE COUNT 92 a 105 c 93 g 94 t  
ORIGIN

Query Match 94.5%; Score 300.4; DB 6; Length 384;  
Best Local Similarity 96.5%; Pred. No. 4.5e-89;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 CAAATGTTCTCTCCAGCTCCAGCAATCTGTGCTCCAGGGGAAAAGGTGACA 60  
DB CAAATGTTCTCTCCAGCTCCAGCAATCTGTGCTCCAGGGGAAAAGGTGACA 126  
61 ATGACTTGAGGGCCAGCTCAAGTGAATTAATCATGCACTGTACAGCAGAGCCAGGA 120  
DB ATGACTTGAGGGCCAGCTCAAGTGAATTAATCATGCACTGTACAGCAGAGCCAGGA 186  
121 TCCTCCCCAAACCTGGATTCTGCACATCAACCTGGCTTGGAGTCCCTGCTGC 180  
DB TCCTCCCCAAACCTGGATTCTGCACATCAACCTGGCTTGGAGTCCCTGCTGC 246

181 TTCAGTGGCAGTGGGCTCTGGACCTTACTCTCTCACAATACGACAGTGGAGCTGAA 240  
DB TTCAGTGGCAGTGGGCTCTGGACCTTACTCTCTCACAATACGACAGTGGAGCTGAA 306

241 GATGCTGCCACTTATTACTGCGAGCTAGTAACCCAGCAGCTTGGAGGGGG 300  
DB GATGCTGCCACTTATTACTGCGAGCTAGTAACCCAGCAGCTTGGAGGGGG 366

301 ACCATGCTGGAATAAGA 318  
DB ACCAAGCTGGAATCAAA 384

RESULT 10  
BD004717  
LOCUS BD004717 384 bp DNA linear PAT 31-JAN-2002  
DEFINITION Chimeric antibody against human B lymphocyte limited differentiation antibody for remedy of B cell lymphoma and therapeutic use of radiolabeled antibody.  
ACCESSION BD004717  
VERSION BD004717.1 GI:18632678  
KEYWORDS JP 2001010974-A/5.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo. Anderson,D.R., Rastetter,W.H., Hanna,N., Leonard,J.E., Newman,R.A. and Reff,M.E.  
TITLE Chimeric antibody against human B lymphocyte limited differentiation antibody for remedy of B cell lymphoma and therapeutic use of radiolabeled antibody  
JOURNAL Patent: JP 2001010974-A 5 16-JAN-2001;  
IDEC PHARMACEUTICALS CORP  
OS Homo sapiens (human)  
PN JP 2001010974-A/5  
PD 16-JAN-2001  
PF 21-APR-2000 JP 2000126317  
PR 13-NOV-1992 US 07/978891.03-NOV-1993 US 08/149099 PI  
DURRELL R ANDERSON,WILLIAM H RASTETTER,NABIL HANNA, PI JOHN E LEONARD,  
PI ROWLAND A NEWMAN,MITCHELL E REFF  
PC A61K39/395,A61K31/375,A61K31/573,A61K31/664,A61K31/704, PC  
A61K45/00,A61P35/00.  
PC C07K16/28,C07K16/46,G01N33/53//C12N5/10,C12N5/09,C12P21/02, PC  
C12P21/08,  
PC (C12P21/02,C12R1:91),C12P21/08,C12R1:91),C12N5/00,C12N5/00 CC  
FH Key  
FT source Location/Qualifiers  
1..384  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 92 a 105 c 93 g 94 t  
ORIGIN

Query Match 94.5%; Score 300.4; DB 6; Length 384;  
Best Local Similarity 96.5%; Pred. No. 4.5e-89;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 CAAATGTTCTCTCCAGCTCCAGCAATCTGTGCTCCAGGGGAAAAGGTGACA 60  
DB CAAATGTTCTCTCCAGCTCCAGCAATCTGTGCTCCAGGGGAAAAGGTGACA 126  
61 ATGACTTGAGGGCCAGCTCAAGTGAATTAATCATGCACTGTACAGCAGAGCCAGGA 120  
DB ATGACTTGAGGGCCAGCTCAAGTGAATTAATCATGCACTGTACAGCAGAGCCAGGA 186  
121 TCCTCCCCAAACCTGGATTCTGCACATCAACCTGGCTTGGAGTCCCTGCTGC 180  
DB TCCTCCCCAAACCTGGATTCTGCACATCAACCTGGCTTGGAGTCCCTGCTGC 246

Db 187 TCCTCCCCCAACCGTGGATTATATCCACATCCAACTGGCTTCTGGAGTCCCTGTTCCG 246  
Qy 181 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 240  
Db 247 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 306  
Qy 241 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 300  
Db 307 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 366  
Qy 301 ACCATGCTGGAATTAAGA 318  
Db 367 ACCAAGCTGGAATTAAGA 384

RESULT 11  
AR000007 9209 bp DNA linear PAT 04-DEC-1998  
LOCUS AR000007  
DEFINITION Sequence 3 from patent US 5736137.  
ACCESSION AR000007  
VERSION AR000007.1 GI:3962538  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9209)  
AUTHORS Anderson,D.R., Hanna,N., Leonard,J.E., Newman,R.A., Reff,M.E. and Rastetter,W.H.  
TITLE Therapeutic application of chimeric and radiolabeled antibodies to human B lymphocyte restricted differentiation antigen for treatment of B cell lymphoma  
JOURNAL Patent: US 5736137-A 3 07-APR-1998;  
FEATURES Location/Qualifiers  
source 1..9209  
BASE COUNT 2239 a 2397 c 2390 g 2183 t  
ORIGIN

Query Match 94.5%; Score 300.4; DB 6; Length 9209;  
Best local Similarity 96.5%; Pred. No. 4.1e-89;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTGTCATCTCCAGGGGAAAGTCCACA 60  
Db 1045 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTGTCATCTCCAGGGGAAAGTCCACA 1104  
Qy 61 ATGACTTGCAGGGGCGCAGCTCAAGTGAATTAATGACATGACATGTCACGAGCAAGCCAGGA 120  
Db 1105 ATGACTTGCAGGGGCGCAGCTCAAGTGAATTAATGACATGACATGTCACGAGCAAGCCAGGA 1164  
Qy 121 TCCCTCCCCCAACCGTGGATTATATCCACATCCAACTGGCTTCTGGAGTCCCTGTTCCG 180  
Db 1165 TCCCTCCCCCAACCGTGGATTATATCCACATCCAACTGGCTTCTGGAGTCCCTGTTCCG 1224  
Qy 181 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 240  
Db 1225 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 1284  
Qy 241 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 300  
Db 1285 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 1344  
Qy 301 ACCATGCTGGAATTAAGA 318  
Db 1345 ACCAAGCTGGAATTAAGA 1362

RESULT 12  
AR060920 9209 bp DNA linear PAT 29-SEP-1999  
LOCUS AR060920  
DEFINITION Sequence 3 from patent US 5843439.  
ACCESSION AR060920  
VERSION AR060920.1 GI:5988611

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9209)  
AUTHORS Anderson,D.R., Hanna,N., Leonard,J.E., Newman,R.A., Reff,M.E. and Rastetter,W.H.  
TITLE Therapeutic application of chimeric and radiolabeled antibodies to human B lymphocyte restricted differentiation antigen for treatment of B cell lymphoma  
JOURNAL Patent: US 5843439-A 3 01-DEC-1998;  
FEATURES Location/Qualifiers  
source 1..9209  
BASE COUNT 2239 a 2397 c 2390 g 2183 t  
ORIGIN

Query Match 94.5%; Score 300.4; DB 6; Length 9209;  
Best local Similarity 96.5%; Pred. No. 4.1e-89;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTGTCATCTCCAGGGGAAAGTCCACA 60  
Db 1045 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTGTCATCTCCAGGGGAAAGTCCACA 1104  
Qy 61 ATGACTTGCAGGGGCGCAGCTCAAGTGAATTAATGACATGACATGTCACGAGCAAGCCAGGA 120  
Db 1105 ATGACTTGCAGGGGCGCAGCTCAAGTGAATTAATGACATGACATGTCACGAGCAAGCCAGGA 1164  
Qy 121 TCCCTCCCCCAACCGTGGATTATATCCACATCCAACTGGCTTCTGGAGTCCCTGTTCCG 180  
Db 1165 TCCCTCCCCCAACCGTGGATTATATCCACATCCAACTGGCTTCTGGAGTCCCTGTTCCG 1224  
Qy 181 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 240  
Db 1225 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 1284  
Qy 241 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 300  
Db 1285 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 1344  
Qy 301 ACCATGCTGGAATTAAGA 318  
Db 1345 ACCAAGCTGGAATTAAGA 1362

RESULT 13  
AR211052 9209 bp DNA linear PAT 20-JUN-2002  
LOCUS AR211052  
DEFINITION Sequence 3 from patent US 6399061.  
ACCESSION AR211052  
VERSION AR211052.1 GI:21514272  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9209)  
AUTHORS Anderson,D.R., Hanna,N., Newman,R.A., Reff,M.E. and Rastetter,W.H.  
TITLE Chimeric and radiolabeled antibodies specific to human CD20 antigen and use thereof for treatment of B-cell lymphoma  
JOURNAL Patent: US 6399061-A 3 04-JUN-2002;  
FEATURES Location/Qualifiers  
source 1..9209  
BASE COUNT 2239 a 2397 c 2390 g 2183 t  
ORIGIN

Query Match 94.5%; Score 300.4; DB 6; Length 9209;  
Best local Similarity 96.5%; Pred. No. 4.1e-89;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTGTCATCTCCAGGGGAAAGTCCACA 60  
Db 1045 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTGTCATCTCCAGGGGAAAGTCCACA 1104  
Qy 61 ATGACTTGCAGGGGCGCAGCTCAAGTGAATTAATGACATGACATGTCACGAGCAAGCCAGGA 120  
Db 1105 ATGACTTGCAGGGGCGCAGCTCAAGTGAATTAATGACATGACATGTCACGAGCAAGCCAGGA 1164  
Qy 121 TCCCTCCCCCAACCGTGGATTATATCCACATCCAACTGGCTTCTGGAGTCCCTGTTCCG 180  
Db 1165 TCCCTCCCCCAACCGTGGATTATATCCACATCCAACTGGCTTCTGGAGTCCCTGTTCCG 1224  
Qy 181 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 240  
Db 1225 TTGAGTGCGAGTGGGTGGGACCTTTACTCTCAGCAATCAGCAGAGTGGAGGCTGAA 1284  
Qy 241 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 300  
Db 1285 GATGCTGCCACTTATTAATCTCCAGCAGTGGAGTAGTAACCCAGCCAGTTCCGAGGGGG 1344  
Qy 301 ACCATGCTGGAATTAAGA 318  
Db 1345 ACCAAGCTGGAATTAAGA 1362



```

Db 1045 CAAATGTTCTCTCCAGCTCTCCAGCAATCTCTGCATCTCCAGGGAGAGAGTCACA 1104
QY 61 ATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGTGTACCAAGCAGAGCAGA 120
Db 1105 ATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGTGTACCAAGCAGAGCAGA 1164
QY 121 TCCCTCCCAAAACCTGGATTTCTCTGCACATCCAGCTGGCTTGGAGTCCCTGCTGC 180
Db 1165 TCCCTCCCAAAACCTGGATTTCTCTGCACATCCAGCTGGCTTGGAGTCCCTGCTGC 1224
QY 181 TTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGAGAGTGGAGCTGAA 240
Db 1225 TTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGAGAGTGGAGCTGAA 1284
QY 241 GATGCTGCCACTTATTACTGCCAGAGTGGAGTAAACCCACCCAGCTTGGAGGGGG 300
Db 1285 GATGCTGCCACTTATTACTGCCAGAGTGGAGTAAACCCACCCAGCTTGGAGGGGG 1344
QY 301 ACCATGCTGGAATAAGA 318
Db 1345 ACCAAGCTGGAAATCAAA 1362

RESULT 14
AX032414 9209 bp DNA circular PAT 20-SEP-2000
LOCUS AX032414
DEFINITION Sequence 3 from Patent EP1005870.
ACCESSION AX032414
VERSION AX032414.1 GI:10279387
KEYWORDS
ORGANISM
SOURCE
REFERENCE
1 (bases 1 to 9209)
AUTHORS
Rastetter,W.H., Hanna,N., Leonard,J.E., Newman,R.A., Reff,M.E. and
Anderson,D.R.
TITLE
Therapeutic application of chimeric antibodies to human b
lymphocyte restricted differentiation antigen for treatment of b
cell lymphoma
JOURNAL
Patent: EP 1005870-A 3 07-JUN-2000;
IDEC PHARMA CORP (US)
FEATURES
Location/Qualifiers:
source 1..9209
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 2239 a 2397 c 2390 g 2183 t
ORIGIN
Query Match 94.5%; Score 300.4; DB 6; Length 9209;
Best Local Similarity 96.5%; Pred. No. 4,1e-89;
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

Db 1345 ACCAAGCTGGAAATCAAA 1362
|||||
RESULT 15
BD004714 9209 bp DNA linear PAT 31-JAN-2002
LOCUS BD004714
DEFINITION
Chimeric antibody against human B lymphocyte limited
differentiation antibody for remedy of B cell lymphoma and
therapeutic use of radiolabeled antibody.
ACCESSION BD004714
VERSION BD004714.1 GI:18632675
KEYWORDS
SOURCE JP 2001010974-A/2.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 9209)
AUTHORS
Anderson,D.R., Rastetter,W.H., Hanna,N., Leonard,J.E., Newman,R.A.
and Reff,M.E.
TITLE
Chimeric antibody against human B lymphocyte limited
differentiation antibody for remedy of B cell lymphoma and
therapeutic use of radiolabeled antibody
Patent: JP 2001010974-A 2 16-JAN-2001;
IDEC PHARMACEUTICALS CORP
COMMENT OS Homo sapiens (human)
PN JP 2001010974-A/2
PD 16-JAN-2001
PE 21-APR-2000 JP 2000126317
PR 13-NOV-1992 US 07/978891, 03-NOV-1993 US 08/149099 PI
DURRELL R ANDERSON, WILLIAM H RASTETTER, NABIL HANNA, PI JOHN E
LEONARD,
PI ROWLAND A NEWMAN, MITCHELL E REFF
PC A6IK39/395, A6IK31/375, A6IK31/573, A6IK31/6C4, A6IK31/704, PC
A6IK45/00, A6IP35/00,
PC C07K16/28, C07K16/46, G01N33/53//C12N5/10, C12N15/09, C12P21/02,
PC C12P21/08,
PC (C12P21/02, C12R1:91), (C12P21/08, C12R1:91), C12N5/00, C12N15/00
CC FC
CH Key
FT source 1..9209
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2239 a 2395 c 2392 g 2183 t
ORIGIN
Query Match 94.5%; Score 300.4; DB 6; Length 9209;
Best Local Similarity 96.5%; Pred. No. 4,1e-89;
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

QY 301 ACCATGCTGGAAATPAGA 318  
||| ||||| |  
Db 1345 ACCAAGCTGGAAATCANA 1362

Search completed: November 27, 2002, 04:41:35  
Job time : 1280.54 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:34:25 : Search time 171.266 Seconds  
(without alignments)  
4181.413 Million cell updates/sec

Title: US-09-893-615-88

Perfect score: 318  
Sequence: 1 CAATGTCTCTCTCCAGTC.....CGACCATCTCTGGAATAGA 318

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300.4	94.5	321	18	AAAT77377
2	300.4	94.5	321	24	ABK23937
3	298.8	94.0	384	15	AAO65630
4	298.4	93.8	19001	19	AAV61793
5	297.2	93.5	449	21	AAA43472
6	297.2	93.5	708	21	AAA63530
7	297.2	93.5	9208	15	AAO65629
8	295.6	93.0	318	18	AAAT79900
9	295.6	93.0	318	24	ABK24005

10	295.6	93.0	335	18	AAAT79899
11	295.6	93.0	335	24	ABK24004
12	295.6	93.0	403	8	AAAT70972
13	295.6	93.0	403	10	AAAT91147
14	295.6	93.0	404	20	AAV82358
15	295.6	93.0	404	22	AAH22071
16	295.6	93.0	426	18	AAAT36317
17	295.6	93.0	426	18	AAAT70869
18	295.6	93.0	426	18	AAAT51043
19	295.6	93.0	426	19	AAV03927
20	295.6	93.0	426	19	AAV18558
21	295.6	93.0	426	19	AAV18594
22	293	92.1	387	13	AAO27350
23	293	92.1	1925	21	AAA15019
24	292.4	91.9	387	13	AAO20983
25	290.8	91.4	732	17	AAAT42508
26	290.8	91.4	732	20	AAV72081
27	289.4	91.0	765	22	AAAC85590
28	289.4	91.0	1239	22	AAAC85563
29	289.4	91.0	1280	22	AAAC85564
30	288.2	90.6	435	12	AAQ15115
31	288.2	90.6	435	14	AAQ36613
32	287.6	90.4	705	20	AAV72047
33	287.6	90.4	3217	20	AAV72076
34	286.6	90.1	318	24	AAAS97128
35	284	89.3	765	22	AAAC85591
36	283.4	89.1	1590	20	AAAX24806
37	281.8	88.6	321	14	AAO39412
38	281.4	88.5	711	15	AAO55180
39	281.4	88.5	711	18	AAAT65006
40	281.2	88.4	758	21	AAA60937
41	281.2	88.4	759	22	AAAC60426
42	280.2	88.1	711	17	AAAT17728
43	280.2	88.1	711	21	AAZ98748
44	278	87.4	321	10	AAAT91661
45	278	87.4	420	12	AAQ11970

#### ALIGNMENTS

RESULT 1	
AAAT77377	
ID	AAAT77377 standard; cDNA; 321 BP.
XX	
AC	AAAT77377;
XX	
DT	26-DEC-1997 (first entry)
XX	
DE	Mouse anti-human Factor IX antibody BC2 light chain cDNA.
XX	
XX	Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;
KW	humanized antibody; antibody engineering; light chain; CDS;
KW	complementarity determining region; myocardial infarction;
KW	angina; atrial fibrillation; stroke; kidney damage;
KW	pulmonary embolism; deep vein thrombosis; coronary angioplasty;
KW	disseminated intravascular coagulation; artificial organ; sepsis;
XX	shunt; prosthesis; ss.
XX	
OS	Mus musculus.
XX	
PN	WO9726010-A1.
XX	
PD	24-JUL-1997.
XX	
PF	17-JAN-1997; 97WO-US00759.
XX	
PR	24-OCT-1996; 96US-0029119.
XX	
PR	17-JAN-1996; 96US-0010108.
XX	
PA	(SMK ) SMITHKLINE BEECHAM CORP.
XX	(UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

Anti-Factor IX Mab  
Murine BC2 light c  
2H7 Vh sequence in  
2H7 Vh sequence.  
Mouse antibody 2H7  
2H7 light chain va  
2H7 antibody light  
2H7 light chain va  
Coding sequence fo  
Mouse 2H7 antibody  
Mouse 2H7 antibody  
Encodes ASB7 antib  
DNA encoding a CD-  
Encodes Variable r  
Murine ASB7 Light  
Plasmid pBE14/ASB7  
DNA encoding a fus  
DNA encoding a fus  
Anti-CD20 single c  
IL-2 chimeric anti  
Anti-IL2R-alpha an  
Plasmid pNG3/ASB7  
IRES-based ASB7 ch  
Anti-RKGD2 hybrid  
DNA encoding a fus  
ASB7 F(ab')2 codin  
Mab32 Light chain  
Sequence encoding  
Single-chain anti-  
Chimeric L6 anti-t  
L6 sFv DNA. Unde  
Anti-erbB2 scFv CD  
Nucleotide sequenc  
Chimeric antibody  
Sequence encoding

PI Blackburn MN, Church WR, Feuerstein GZ, Gross MS;  
PI Nichols AJ, Padlan EA, Patel AH, Sylvester DR;  
XX  
DR WPI: 1997-385117/35.  
P-PSDB: AAMW24520.  
XX  
PT Inhibiting thrombosis with self-limiting antibody to coagulation  
PT factor - avoids uncontrolled bleeding by providing only partial  
PT inhibition  
XX  
PS Example 5; Page 64; 150pp; English.  
XX  
CC This cDNA sequence encodes the light chain variable region (see  
CC AAMW24520) of mouse anti-human factor IX monoclonal antibody BC2.  
CC Claimed humanised antibodies (see AAMW24510-18) contain CDRs (see  
CC AAMW24504-09) of BC2 heavy and light chains inserted into framework  
CC regions of selected human antibody sequences. They have self-  
CC limiting neutralising activity, and are useful as anticoagulant  
CC agents in treatment of thrombosis associated with myocardial  
CC infarction, unstable angina, atrial fibrillation, stroke, renal  
CC damage, pulmonary embolism, deep vein thrombosis, percutaneous  
CC transluminal coronary angioplasty, disseminated intravascular  
CC coagulation, sepsis, or artificial organs, shunts or prostheses  
CC (claimed). Also claimed are chimeric antibodies (see AAT79900), and  
CC Fab and Fab')2 fragments. The claimed antibodies do not cause  
CC uncontrolled bleeding (contrast heparin and warfarin) since they  
CC provide only partial inhibition of coagulation.  
XX  
SQ Sequence 321 BP; 78 A; 92 C; 79 G; 72 T; 0 other;  
XX  
Query Match 94.5%; Score 300.4; DB 18; Length 321;  
Best Local Similarity 96.5%; Pred. No. 3.3e-77;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
XX  
QY 1 CAAATTTGTCCTCCAGCTCCAGCAATCTGTCGATCCAGGGGAAAGGTACA 60  
Db 1 CAAATTTGTCCTCCAGCTCCAGCAATCTGTCGATCCAGGGGAGAGGTACA 60  
QY 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 120  
Db 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 120  
QY 121 TCCCTCCCAAAACCTGGATTATGCAATCCAACTGGCTTCGGAGTCCCTGCTGC 180  
Db 121 TCCCTCCCAAAACCTGGATTATGCAATCCAACTGGCTTCGGAGTCCCTGCTGC 180  
QY 181 TTCACTGGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 240  
Db 181 TTCACTGGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 240  
QY 241 GATGCTGCCAATTTACTGTCACAGTGAAGTAAACCAAGTTCGGAGGGGG 300  
Db 241 GATGCTGCCAATTTACTGTCACAGTGAAGTAAACCAAGTTCGGAGGGGG 300  
QY 301 ACCATGCTGGAATTAAGA 318  
Db 301 ACCAAGCTGGAATTAAGA 318  
XX  
RESULT 2  
ABK23937  
ID ABK23937 standard; cDNA; 321 BP.  
XX  
AC ABK23937;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Murine BC2 light chain variable region cDNA.  
XX  
KM Human: mouse; BC2: animal post-thromboembolic induced ischemia;  
KM thrombolytic agent; anti-factor IX antibody; plasminogen activator; ss;  
KM thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;  
KM vasotropic; cardiant; PCR primer; anti-respiratory syncytial virus;  
XX

KW heavy chain variable region; light chain variable region.  
XX  
OS Mus sp.  
XX  
PN WO200187339-A1.  
XX  
PD 22-NOV-2001.  
XX  
PF 05-OCT-2000; 2000MO-US27438.  
XX  
PR 15-MAY-2000; 2000US-0571434.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
XX  
PI Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;  
XX  
DR WPI: 2002-082944/11.  
P-PSDB: AAU80976.  
XX  
PT Treating post-thromboembolic induced ischemia in an animal by  
PT administering anti-factor IX antibody in combination with a plasminogen  
PT activator -  
XX  
PS Example 5; Page 94; 163pp; English.  
XX  
CC The invention relates to a method for treating an animal  
CC post-thromboembolic induced ischemia or reducing a required dose of a  
CC thrombolytic agent in treatment of an animal post-thromboembolic induced  
CC ischemia, comprising administering an anti-factor IX antibody or its  
CC fragment, optionally in combination with a plasminogen activator or  
CC thrombolytic agent. The method is useful for treating thromboembolic  
CC post-thromboembolic-induced ischemia, for preventing thromboembolic  
CC stroke in an animal, and for reducing a required dose of a thrombolytic  
CC agent. Sequences ABK23932-ABK24009 represent DNA molecules encoding  
CC antibodies and PCR primers used in the method of the invention.  
XX  
SQ Sequence 321 BP; 78 A; 92 C; 79 G; 72 T; 0 other;  
XX  
Query Match 94.5%; Score 300.4; DB 24; Length 321;  
Best Local Similarity 96.5%; Pred. No. 3.3e-77;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
XX  
QY 1 CAAATTTGTCCTCCAGCTCCAGCAATCTGTCGATCCAGGGGAAAGGTACA 60  
Db 1 CAAATTTGTCCTCCAGCTCCAGCAATCTGTCGATCCAGGGGAGAGGTACA 60  
QY 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 120  
Db 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 120  
QY 121 TCCCTCCCAAAACCTGGATTATGCAATCCAACTGGCTTCGGAGTCCCTGCTGC 180  
Db 121 TCCCTCCCAAAACCTGGATTATGCAATCCAACTGGCTTCGGAGTCCCTGCTGC 180  
QY 181 TTCACTGGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 240  
Db 181 TTCACTGGACAGGGCCAGCTCAAGTAAATTACATGCATGTCACAGACCAAGCA 240  
QY 241 GATGCTGCCAATTTACTGTCACAGTGAAGTAAACCAAGTTCGGAGGGGG 300  
Db 241 GATGCTGCCAATTTACTGTCACAGTGAAGTAAACCAAGTTCGGAGGGGG 300  
QY 301 ACCATGCTGGAATTAAGA 318  
Db 301 ACCAAGCTGGAATTAAGA 318  
XX  
RESULT 3  
AAO65630  
ID AAO65630 standard; DNA; 384 BP.  
XX  
AC AAO65630;  
XX

DT	07-JUN-1999	(first entry)
XX		
XX	target plasmid	Molly containing anti-CD20 gene.
XX		
KW	Molly: target plasmid; gene integration; gene amplification;	
KW	homologous recombination; vector: neomycin phosphotransferase;	
KW	neo gene; selectable marker; immunoglobulin; CD20; C288; human; s	
XX		
OS	Chimeric - Mus sp.	
OS	Chimeric - Escherichia coli.	
OS	Chimeric - Baculovirus.	
OS	Chimeric - Cytomegalovirus.	
OS	Chimeric - Rhesus macaque polyoma virus.	
OS	Chimeric - Pottius sp.	
OS	Chimeric - Salmonella typhimurium.	
OS	Chimeric - Homo sapiens.	
XX		
FM	Key	Location/Qualifiers
FM	misc_feature	361..363
FM		/tag= a
FM		/note= "these bases represent nucleotides missing from the sequence given in the specification. They are included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	721..722
FM		/tag= b
FM		/note= "these bases represent nucleotides missing from the sequence given in the specification. They are included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	2942
FM		/tag= c
FM		/note= "this base represents a nucleotide missing from the sequence given in the specification. It is included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	3301
FM		/tag= d
FM		/note= "this base represents a nucleotide missing from the sequence given in the specification. It is included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	4261..4262
FM		/tag= e
FM		/note= "these bases represent nucleotides missing from the sequence given in the specification. They are included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	4621..4622
FM		/tag= f
FM		/note= "these bases represent nucleotides missing from the sequence given in the specification. They are included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	5581
FM		/tag= g
FM		/note= "this base represents a nucleotide missing from the sequence given in the specification. It is included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	5941
FM		/tag= h
FM		/note= "this base represents a nucleotide missing from the sequence given in the specification. It is included to maintain the nucleotide numbering in the specification for this sequence"
FM	misc_feature	5941
FM		/tag= h
FM		/note= "this base represents a nucleotide missing from the sequence given in the specification. It is included to maintain the nucleotide numbering in the specification for this sequence"

```
FT misc_feature 7201 specification for this sequence"
FT /*tag= 1
FT /*note= "this base represents a nucleotide missing
FT from the sequence given in the
FT specification. It is included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 8161..8162
FT /*tag= j
FT /*note= "these bases represent nucleotides missing
FT from the sequence given in the
FT specification. They are included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 8521..8522
FT /*tag= k
FT /*note= "these bases represent nucleotides missing
FT from the sequence given in the
FT specification. They are included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 11101
FT /*tag= l
FT /*note= "this base represents a nucleotide missing
FT from the sequence given in the
FT specification. It is included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 12061..12062
FT /*tag= m
FT /*note= "these bases represent nucleotides missing
FT from the sequence given in the
FT specification. They are included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 12421..12422
FT /*tag= n
FT /*note= "these bases represent nucleotides missing
FT from the sequence given in the
FT specification. They are included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 13381..13382
FT /*tag= o
FT /*note= "these bases represent nucleotides missing
FT from the sequence given in the
FT specification. They are included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 15961..15962
FT /*tag= p
FT /*note= "these bases represent nucleotides missing
FT from the sequence given in the
FT specification. They are included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 16321
FT /*tag= q
FT /*note= "this base represents a nucleotide missing
FT from the sequence given in the
FT specification. It is included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 18541
FT /*tag= r
FT /*note= "this base represents a nucleotide missing
FT from the sequence given in the
FT specification. It is included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT misc_feature 18901
FT /*tag= s

FT /*note= "this base represents a nucleotide missing
FT from the sequence given in the
FT specification. It is included to
FT maintain the nucleotide numbering in the
FT specification for this sequence"
FT W09841645-A1.
FT PD 24-SEP-1998.
FT PF 09-MAR-1998; 98WO-US03935.
FT PR 13-FEB-1998; 98US-0023715.
FT PR 14-MAR-1997; 97US-0819866.
FT PA (IDEC PHARM CORP.
FT PI Barnett RS, McLachlan KR, Refl ME;
FT DR WPI; 1998-521229/44.
FT XX
FT XX Site specific integration of DNA in mammals for expressing, e.g.
FT PT immunoglobulins - comprises homologous recombination using
FT PT selectable marker and target plasmids.
FT PS Example 1; Fig 8; 114pp: English.
FT XX
FT XX This is the nucleotide sequence of novel target plasmid Molly.
FT CC The plasmid includes the murine dihydrofolate reductase (DHFR)
FT CC gene, the Escherichia coli beta-galactosidase gene, baculovirus
FT CC DNA, a cassette comprising the promoter and enhancer elements from
FT CC cytomegalovirus and SV40 virus, the E. coli beta-glucuronidase
FT CC (GUS) gene, firefly luciferase gene, salmonella typhimurium
FT CC histidinol dehydrogenase gene (HisD) and transposon Tn5 neomycin
FT CC phosphotransferase (neo) gene sequences, in a pBR-derived backbone,
FT CC and also an anti-B cell antigen CD20 chimeric antibody C2B8 gene.
FT CC The invention provides a novel method for integrating a desired
FT CC exogenous DNA at a target site within the genome of a mammalian
FT CC cell via homologous recombination. This involves transfecting the
FT CC cell with a 'marker plasmid' such as Desmond (see AAV61792), which
FT CC contains a unique sequence that is foreign to the mammalian cell
FT CC genome and which provides a substrate for homologous recombination,
FT CC followed by transfection with a 'target plasmid', such as Molly
FT CC or Mandy (see AAV61794), containing a sequence which provides for
FT CC homologous recombination with the unique sequences contained in
FT CC the marker plasmid, and further comprising a desired DNA that is
FT CC to be integrated into the mammalian cells, typically an
FT CC immunoglobulin or other secreted mammalian glycoprotein. The
FT CC homologous recombination system utilizes the neo gene as a
FT CC dominant selectable marker. The neo gene is split into 3 exons.

Query Match Score 93.8%; DB 19; Length 19001;
Best local similarity 95.9%; Pred. No. 3,8e-76;
Matches 305; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGTCCTCAGCAATCTCTGTCATCTCCAGGGAAGGTCACA 60
DB 8070 CAAATGTTCTCTCCAGTCCTCAGCAATCTCTGTCATCTCCAGGGAAGGTCACA 8129
QY 61 ATGACTTGAGGGGCGAGCTCAAGTAAATTACATGACGTGAGCAGCAGAGCCAGGA 120
DB 8130 ATGACTTGAGGGGCGAGCTCAAGTAAATTACATGACGTGAGCAGCAGAGCCAGGA 8189
QY 121 TCCTCCGCCCAACCTGATTTCTGACACATCCACCTGCTTCTGAGTCCCTGCTCGC 180
DB 8190 TCCTCCGCCCAACCTGATTTCTGACACATCCACCTGCTTCTGAGTCCCTGCTCGC 8249
QY 181 TTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGCAGAGTGGAGCTGAA 240
DB 8250 TTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGCAGAGTGGAGCTGAA 8309
QY 241 GATGCTGCCACTTATCTAGTCCGACGTGAGTAGTAACCCACCGCTGGAGGGGG 300
|||||
```

DB	8310	GATGCTGCCACATTATTACGCGCAGCAGTGAGTACGTAACCCACCACCGTTCCGAGGGGGG	8366
QY	301	ACCATGCTGGAAATAAGA	318
Db	8370	ACCAAGCTGGAAATCAAA	8387
RESULT 5			
XX	AAAA3472	standard: CDNA: 449 BP.	
XX	AAAA3472;		
XX	21-AUG-2000	(first entry)	
XX			
DE		Mouse secreted expressed sequence tag SEQ ID NO:47.	
KM		Human: mouse; chicken; rat; secreted expressed sequence tag; SEST;	
KM		expressed sequence tag; EST; probe: chemotactic; proliferative;	
KM		immunomodulatory; hematopoietic; chemokinetic; analgesic; hemostatic	
KM		thrombolytic; antinflammatory; cyostatic; antibacterial; antifungal;	
KM		antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;	
KM		antilicer; osteoprotective; neuroprotective; nocotropic; antipsoriatic;	
KM		cerebroprotective; anticonvulsant; antidepressant; gene therapy;	
KM		vaccine; autoimmune disorder; multiple sclerosis; allergic condition;	
KM		insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;	
KM		lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;	
KM		central nervous system disorder; Alzheimer's disease; stroke;	
KM		parkinson's disease; Huntington's disease; coagulation disorder;	
KM		haemophilia; thrombosis; inflammatory disorder; Crohn's disease;	
KM		tumour; infection; depression; psoriasis; ss.	
XX			
OS		Mus musculus.	
XX			
XX		WO200021991-A1.	
XX			
PD	20-APR-2000.		
XX			
PF	15-OCT-1999;	99WO-US24206.	
XX			
PR	15-OCT-1998;	98US-0104436.	
XX			
PA		(GEMV ) GENETICS INST INC.	
XX			
PI	Jacobs K, McCoy JM, Lavallee ER, Collins-Racle LA, Evans C;		
PI	Merberg D, Treacy M, Bowman MR;		
DR			
XX	WPI: 2000-317938/27.		
XX			
PT			
PT	Isolated polynucleotides, and encoded proteins, comprising secreted		
PT	expressed sequence tags (SESTs), useful for treating various disorders		
XX	such as autoimmune, infectious, and central nervous system disorders -		
XX			
PS	Claim 1; Page 214; 803pp: English.		
XX			
XX	AAA43426 to AAA45925 represent specifically claimed secreted expressed		
CC	sequence tags (SESTs), isolated from human, mouse, chicken and rat		
CC	tissue sources. The SESTs can have a range of activities depending on		
CC	the tissues they were isolated from. The activities include:		
CC	chemotactic; proliferative; immunomodulatory; hematopoietic;		
CC	chemokinetic; analgesic; hemostatic; thrombolytic; antinflammatory;		
CC	cytostatic; antibacterial; antifungal; antiviral; antidiabetic;		
CC	antiasthmatic; vulnery; antilicer; osteopathic; neuroprotective;		
CC	nocotropic; antiparkinsonian; antipsoriatic; cerebroprotective;		
CC	anticonvulsant; and antidepressant. The SESTs can be used for gene		
CC	therapy and in vaccines. The SESTs are useful as probes for the		
CC	identification and isolation of full-length cDNAs and genomic DNA		
CC	molecules which correspond to the SESTs. Proteins encoded by the SESTs		
CC	are useful in assays for determining biological activity and raising		
CC	antibodies. They may be useful for treatment of autoimmune disorders		
CC	(multiple sclerosis, insulin dependent diabetes), allergic conditions		
CC	(asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers		
CC	osteoporosis, osteoarthritis, central nervous system disorders		

CC	(Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's disease), tumours, bacterial, fungal or viral infections, depression and psoriasis. AAA45926 to AAA45931 represent linker variants which are given in the exemplification of the present invention.
XX	
S0	Sequence 449 BP; 110 A; 124 C; 107 G; 108 T; 0 other:
	Query Match 93.5%; Score 297.2; DB 21; Length 449;
	Best Local Similarity 95.9%; Pred. No. 3.1e-76;
	Matches 305; Conservative 0; Mismatches 13; Indels 0; Gaps 0
Oy	1 CAATTTGTTCTTCCCGAGCTCCAGCAATTCCTCTGCATCTCCACGGGAAAAGTCA 60
Db	90 CAAATTGTTCTTCTCCAGTCTCCAGCAATCTCTGTGCATCTCCAGGGAGAGGTCA 149
Oy	61 ATGACTTGACGGGCGACACTCAAGTGTAAATCATGTCAGTGTACCAGACAGACAGA 120
Db	150 ATGACTTGACGGGCGACACTCAAGTGTAAATCATGTCAGTGTACCAGACAGACAGA 209
Oy	121 TCTCTCCCCAAACCTGGATTTTGTGCACATCCAACTGGCTTTGAGATCCCTGCTGC 180
Db	210 TCTCTCCCCAAACCTGGATTTATGCAATCCAACTGGCTTTGAGATCCCTGCTGC 269
Oy	181 TTTCAGTGGCAGTGGGCTGTGGGACCTCTTACTCTCTCACAATACAGAGAGTGA 240
Db	270 TTTCAGTGGCAGTGGGCTGTGGGACCTCTTACTCTCTCACAATACAGAGAGTGA 329
Oy	241 GATGCTGCACACTTATTACTGCCAGCAGTGGAGTAGTAACCCACGCTTGGAGGGGG 300
Db	330 GATGCTGCACACTTATTACTGCCAGCAGTGGAGTAGTAACCCGTTGGAGGGGG 389
Oy	301 ACCATGCTGGAATACGA 318
Db	390 ACCAAGCTGGAAATCAAA 407
RESULT 6	
AAA63530	
ID	AAA63530 standard; DNA; 708 BP.
XX	
AC	AAA63530;
XX	
DT	14-NOV-2000 (first entry)
XX	
DE	DNA encoding a dimeric anti-CD20 light chain polypeptide.
XX	
KW	Anti-CD20 antibody; dimeric immunoglobulin; immunoglobulin; IgG;
KM	complement system; Fc gamma receptor; cytotoxic effector cell;
KW	host immune cell; programmed cell death; allergic disorder; cancer;
KW	autoimmune disease; allergic asthma; atopic dermatitis; Crohn's disease;
KW	allergic bronchopulmonary aspergillosis; allergic rhinitis;
KW	Graves's disease; food allergy; allergic contact dermatitis; cancer;
KW	B-cell lymphoma; rheumatoid arthritis; ulcerative colitis; psoriasis;
KW	pigeon breeder's disease; hepatitis; leprosy; Lyme disease;
KW	diabetes mellitus; candidiasis; aplastic anaemia; ss.
XX	
OS	Chimeric - Mus sp.
OS	Chimeric - Homo sapiens.
XX	
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS 1..708
FT	/*tag- a
FT	/product= "dimeric anti-CD20 light chain polypeptide"
FT	sig_peptide 1..66
FT	/*tag- b
FT	misc_feature 67..384
FT	/*tag- c
FT	/note= "encodes murine anti-human CD20 variable light chain"
FT	misc_feature 385..708
FT	/*tag- d
FT	/note= "encodes human kappa light chain constant region"

XX WO20004788-A1.  
XX  
XX 03-AUG-2000.  
XX  
XX 28-JAN-2000; 2000WO-US01893.  
XX  
XX 28-JAN-1999; 99US-0238741.  
XX  
XX (IDEC-) IDEC PHARM CORP.  
XX  
XX Braslowsky GR, Hanna N, Harlharan K, Labarre MJ, Huynh TB;  
XX  
XX WPI: 2000-514811/46.  
XX  
XX P-PSDB: AAB08025.  
XX  
XX Genetically engineering immunoglobulin (Ig) G/IgG dimers for the  
XX treatment of cancers, allergic disorders and autoimmune conditions -  
XX  
XX Example 1; Fig 1A-B; 65pp; English.

XX The present sequence encodes a dimeric anti-CD20 light chain  
XX polypeptide. The dimeric immunoglobulin is used in the method of the  
XX invention. The specification describes a method for producing an  
XX immunoglobulin (Ig) G/IgG dimer. The method comprises genetically  
XX engineering a monoclonal antibody to introduce a cysteine molecule  
XX which inhibits formation of intramolecular disulphide bridges between  
XX sister heavy chains on the same antibody molecule. The dimer is a  
XX homodimer or heterodimer that is capable of activating components of the  
XX complement system, and has the ability to activate and kill cells via the  
XX complement cascade. The dimer is also capable of binding to Fc gamma  
XX receptors on cytotoxic effector cells and on host immune cells, and is  
XX capable of initiating programmed cell death. The IgG/IgG dimers may be  
XX used to treat allergic disorders, cancers and autoimmune diseases such  
XX as allergic asthma, allergic bronchopulmonary aspergillosis, allergic  
XX rhinitis, atopic dermatitis, Crohn's disease, Graves's disease, food  
XX allergies, allergic contact dermatitis, CLL cancers and/or B-cell  
XX lymphomas. They may also be used to treat a range of other diseases and  
XX disorders such as rheumatoid arthritis, ulcerative colitis, psoriasis,  
XX Crohn's disease, hepatitis, leprosy, Lyme disease, diabetes  
XX mellitus, candidiasis and aplastic anaemia. They are also useful for  
XX inducing hyper-cross-linking of membrane antigens and for the  
XX preferential killing of selected cell populations.

XX Sequence 708 BP; 180 A; 196 C; 178 G; 154 T; 0 other;

XX Query Match 93.5%; Score 297.2; DB 21; Length 708;  
XX Best Local Similarity 95.9%; Pred. No. 3.5e-76;  
XX Matches 305; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAGTTCACA 60  
DB 67 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAGTTCACA 126  
QY 61 ATGACTTGGAGGGGCGAGCTCACTGAATTAATTCATGCTGTTACCAAGCAAGCA 120  
DB 127 ATGACTTGGAGGGGCGAGCTCACTGAATTAATTCATGCTGTTACCAAGCAAGCA 186  
QY 121 TCCTCCCGCCAAACCGATTTCTGCAATCCAACTGGCTTCTGGAGTCCCTGCTGCG 180  
DB 187 TCCTCCCGCCAAACCGATTTCTGCAATCCAACTGGCTTCTGGAGTCCCTGCTGCG 246  
QY 181 TTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGACAGAGTGGAGCTGAA 240  
DB 247 TTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGACAGAGTGGAGCTGAA 306  
QY 241 GATGCTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGACAGAGTGGAGCTGAA 300  
DB 307 GATGCTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGACAGAGTGGAGCTGAA 366  
QY 301 ACCATGCTGGAAATTAAGA 318  
DB 367 GCCAAGCTGGAAATTAAGA 384

RESULT 7  
AA065629  
ID AA065629 standard; DNA; 9208 BP.

XX AA065629;  
XX  
XX 01-FEB-1995 (first entry)  
XX  
XX Vector contg. TCAE 8 DNA.

XX B cell lymphoma chimeric antibody: CD20; peripheral blood cells;  
XX cell lysis; ss.  
XX  
XX Synthetic.

XX WO9411026-A.

XX 26-MAY-1994.

XX 12-NOV-1993; 93WO-US10953.

XX 13-NOV-1992; 92US-0978891.  
XX 03-NOV-1993; 93US-0149099.

XX (IDEC-) IDEC PHARM CORP.

XX Anderson DR, Hanna N, Leonard JE, Newman RA, Rastetter WH;  
XX Refine;  
XX  
XX WPI: 1994-183162/22.

XX Treating B cell lymphoma with chimeric antibody - against CD20,  
XX causing rapid depletion of peripheral B cells, also new  
XX antibodies and hybridomas

XX Disclosure; Fig 3; 101pp; English.

XX The sequence shows a vector contg. TCAE8, a gene encoding a chimeric  
XX anti-CD20 antibody for treatment of B cell lymphomas. TCAE8  
XX contains 4 transcriptional cassettes; human Ig light and heavy chain  
XX constant regions, dihydrofolate reductase, neomycin phosphotransferase  
XX and murine variable regions. The vector can be used to produce  
XX antibodies which cause depletion of peripheral blood B cells,  
XX including those associated with lymphoma. They mediate complement-  
XX dependent lysis and lyse target cells by antibody-dependent cellular  
XX cytotoxicity.  
XX See also AA065629-35.

XX Sequence 9208 BP; 2237 A; 2399 C; 2388 G; 2182 T; 2 other;

XX Query Match 93.5%; Score 297.2; DB 15; Length 9208;  
XX Best Local Similarity 95.9%; Pred. No. 7e-76;  
XX Matches 305; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAGTTCACA 60  
DB 1044 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAGTTCACA 1103  
QY 61 ATGACTTGGAGGGGCGAGCTCACTGAATTAATTCATGCTGTTACCAAGCAAGCA 120  
DB 1104 ATGACTTGGAGGGGCGAGCTCACTGAATTAATTCATGCTGTTACCAAGCAAGCA 1163  
QY 121 TCCTCCCGCCAAACCGATTTCTGCAATCCAACTGGCTTCTGGAGTCCCTGCTGCG 180  
DB 1164 TCCTCCCGCCAAACCGATTTCTGCAATCCAACTGGCTTCTGGAGTCCCTGCTGCG 1223  
QY 181 TTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGACAGAGTGGAGCTGAA 240  
DB 1224 TTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGACAGAGTGGAGCTGAA 1283  
QY 241 GATGCTGGCAGTGGGTCTGGGACCTTCTACTCTCTCAATCAGACAGAGTGGAGCTGAA 300



DB 1284 GATGCTGCCACTTATTACTGCGAGGTGACTAGTAACCAACCAGCTTGGAGGGGG 1343

OY 301 ACCATGCTGGAATAAGA 318

DB 1344 ACCAAGCTGGAATCAAA 1361

#### RESULT 8

AAT79900 standard; cDNA, 318 BP.

AC AAT79900;

DT 27-DEC-1997 (first entry)

DE Anti-Factor IX Mab chimeric light chain cDNA.

KW Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;

KW chimeric antibody; antibody engineering; light chain; ss.

OS Chimeric Mus musculus.

OS Chimeric Homo sapiens.

OS Chimeric synthetic.

PN WO9726010-A1.

PD 24-JUL-1997.

PF 17-JAN-1997; 97WO-US00759.

PR 24-OCT-1996; 96US-0029119.

PR 17-JAN-1996; 96US-0010108.

PA (SMK) SMITHKLINE BEECHAM CORP.

PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

PI Blackburn MN, Church WR, Feuerstein GZ, Gross MS,

PI Nichols AJ, Padlan EA, Patel AH, Sylvester DR.

DR WPI: 1997-385117/35.

DR P-PSDB: AAW24532.

XX Inhibiting thrombosis with self-limiting antibody to coagulation

PT factor - avoids uncontrolled bleeding by providing only partial

PT inhibition

PS Example 7; Page 128; 150pp; English.

XX This cDNA sequence encodes a mouse-human chimeric antibody

CC light chain (AAW24532) in which the variable region is derived

CC from mouse anti-human factor IX monoclonal antibody BC2 cDNA (see

CC AAT79899) and human sequences from the immunoglobulin Rf-TS3/CL

CC framework. It was obtained by PCR amplification (see AAT79897-98)

CC of BC2 cDNA and insertion of the PCR product into p9HHC 1-3 cDNA

CC (see AAT77374). Claimed anti-factor IX chimeric antibodies are

CC useful in the treatment of thrombosis.

XX Sequence 318 BP; 79 A; 91 C; 78 G; 70 T; 0 other;

XX Query Match 93.0%; Score 295.6; DB 18; Length 318;

XX Best Local Similarity 95.6%; Pred. No. 8.1e-76;

XX Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 CAAATGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAGGTCA 60

DB 1 CAGATAGTACTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAGGTCA 60

OY 61 ATGACTTGCAGGGCCAGCTCAGTGTAAATTCATGCACTGTACCAAGAGCCAGA 120

DB 61 ATGACTTGCAGGGCCAGCTCAGTGTAAATTCATGCACTGTACCAAGAGCCAGA 120

OY 121 TCCTCCCAACCCCTGATTTCTGCGACATCCACACTGCTTCTGAGAGTCCCTGCTCC 180

DB 121 TCCTCCCAACCCCTGATTTATGCGACATCCACACTGCTTCTGAGAGTCCCTGCTCC 180

OY 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTTCACATCAGCAGTGGAGCTGAA 240

DB 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTTCACATCAGCAGTGGAGCTGAA 240

OY 241 GATGCTGCCACTTATTACTGCGAGGTGACTAGTAACCAACCAGCTTGGAGGGGG 300

DB 241 GATGCTGCCACTTATTACTGCGAGGTGACTAGTAACCAACCAGCTTGGAGGGGG 300

OY 301 ACCATGCTGGAATAAGA 318

DB 301 ACCAAGCTGGAATCAAA 318

#### RESULT 9

ABK24005 standard; cDNA, 318 BP.

AC ABK24005;

DT 09-APR-2002 (first entry)

DE Mouse-human light chain DNA.

KW Human; mouse; BC2; animal post-thromboembolic induced ischaemia;

KW thrombolytic agent; anti-factor IX antibody; plasminogen activator; ss;

KW thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;

KW vasotropic; cardiatic; PCR primer; anti-respiratory syncytial virus;

XX heavy chain variable region; light chain variable region.

OS Chimeric - Mus sp.

OS Chimeric - Homo sapiens.

PN WO200187339-A1.

PD 22-NOV-2001.

PF 05-OCT-2000; 2000WO-US27438.

PR 15-MAY-2000; 2000US-0571434.

PA (SMK) SMITHKLINE BEECHAM CORP.

PA Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;

DR WPI: 2002-082944/11.

DR P-PSDB: AAN81002.

XX Treating post-thromboembolic induced ischaemia in an animal by

PT administering anti-factor IX antibody in combination with a plasminogen

PT activator

PS Example 7; Page 154-155; 163pp; English.

XX The invention relates to a method for treating an animal

CC post-thromboembolic induced ischaemia or reducing a required dose of a

CC thrombolytic agent in treatment of an animal post-thromboembolic induced

CC ischaemia, comprising administering an anti-factor IX antibody or its

CC fragment, optionally in combination with a plasminogen activator or

CC thrombolytic agent. The method is useful for treating thromboembolic

CC post-thromboembolic-induced ischaemia, for preventing thromboembolic

CC stroke in an animal, and for reducing a required dose of a thrombolytic

CC agent. Sequences ABK23932-ABK24009 represent DNA molecules encoding

CC antibodies and PCR primers used in the method of the invention.

XX Sequence 318 BP; 79 A; 91 C; 78 G; 70 T; 0 other;

XX Query Match 93.0%; Score 295.6; DB 24; Length 318;

XX Best Local Similarity 95.6%; Pred. No. 8.1e-76;

XX Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy	1	CANAATGTTCTTCCTCCCAATCTCCAGCAAAATCTGCTGATGTCCAGGGGAAAAGGTACA	60
Oy	1		60
Db	1	CAGATGAGTACTCTCCCACTCTCCAGCAATCTGCTGATGTCCAGGGGAAAAGGTACA	60
Oy	61	ATGACTTTCAGAGGGCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAAGACAAGCCAGGA	120
Db	61	ATGACTTTCAGAGGGCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAAGACAAGCCAGGA	120
Oy	121	TGCTCCCCCAAAACCCGTGATTTTGCCCAATCCAACTGGCTTGTGAGTCCCTGCTCGC	180
Db	121	TGCTCCCCCAAAACCCGTGATTTATGAGCCATCCAACTGGCTTGTGAGTCCCTGCTCGC	180
Oy	181	TTTCAGTGGCAGAGGGGTGGGACCTCTACTCTCTCAAAATCAGACAGAGTGGAGGCTGAA	240
Db	181	TTTCAGTGGCAGAGGGGTGGGACCTCTACTCTCTCAAAATCAGACAGAGTGGAGGCTGAA	240
Oy	241	GATGCTGGCAGTATTTACTGGCAGAGAGTGTAGTAACCAACCCAGCTTTCGAGAGGGGG	300
Db	241	GATGCTGGCAGTATTTACTGGCAGAGAGTGTAGTAACCAACCCAGAGGAGTTTCGAGAGGC	300
Oy	301	ACCATGCTGGAAATTAAGA 318	
Db	301	ACCAAGCTGGAAATTAAGA 318	
RESULT 10			
ID	AAT79899	standard; cDNA: 335 BP.	
AC	AAT79899;		
DT	27-DEC-1997	(first entry)	
DE	Anti-Factor IX MAB BC2 light chain PCR product.		
XX			
KM	Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;		
KM	chimeric antibody; antibody engineering; light chain; ss.		
XX			
OS	Chimeric Mus musculus.		
OS	Chimeric synthetic.		
PN	W09726010-A1.		
PD	24-JUL-1997.		
XX			
PF	17-JAN-1997; 97WO-US00759.		
XX			
PR	24-OCT-1996; 96US-0029119.		
XX			
PR	17-JAN-1996; 96US-0010108.		
XX			
PA	(SMIK ) SMITHKLINE BEECHAM CORP.		
PA	(UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.		
PI	Blackburn MN, Church WR, Feuerstein GZ, Gross MS;		
PI	Nichols AJ, Padian EA, Patel AH, Sylvester DR;		
XX			
DR	WPI: 1997-385117/35.		
DR	P-PSDB: AAW24531.		
PT	Inhibiting thrombosis with self-limiting antibody to coagulation		
PT	factor - avoids uncontrolled bleeding by providing only partial		
PT	inhibition		
XX			
PS	Example 7: Page 126; 150bp; English.		
XX			
CC	This cDNA sequence was obtained by PCR amplification (see AAT79897		
CC	and AAT79998) of the light chain variable region (see also AAT77377)		
CC	of mouse anti-human factor IX monoclonal antibody BC2 cDNA. The		
CC	amplification resulted in the addition of ScaI, Nari ends to the		
CC	VL region. The PCR product was ligated into ScaI, Nari-digested		
CC	pFgZHC 1-3 (see AAT77374) and digested with ScaI, Nari to produce a		
CC	mouse-human chimeric light chain FgCHLC (see AAT79900, AAW24532).		
CC	mouse anti-Factor IX chimeric antibodies are useful in the		

CC	treatment of thrombosis.
XX	
SQ	Sequence 335 BP; 80 A; 97 C; 85 G; 73 T; 0 other;
	Query Match            93.0%; Score 295.6; DB 18; Length 335;
	Best Local Similarity 95.6%; Pred. No. 8.2e-76;
	Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY	1 CAATTTGTTCTCTGCCAGTCTGCACCAATTCGTCTGCATCTCAGGGCGAAAAGGTACA 60
DB	1 CAGATAGTACTCTCCAGTCTCCAGCATTCACCAATTCCTGTGCATCTCCAGGGAGAAAGTACACA 60
OY	61 ATGACTTCGAGGGCGAGCTCAAGTGTAATTATCATGCACTGGTNCCACACAGAACGCACAGA 120
DB	61 ATGACTTCGAGGGCGAGCTCAAGTGTAATTATCATGCACTGGTACCACACAGAACGCACAGA 120
OY	121 TCCTCCCCCAAACCTGGATTTCTGCCAATCCAACTGGCTTCTGGAGTCCCTGCTGC 180
DB	121 TCCTCCCCCAAACCTGGATTTATGTCACATCCAACTGGCTTCTGGAGTCCCTGCTGC 180
OY	181 TTCACTGGCAGTGGGTCTGGGACCTTACTCTCTCACAAATCAGCAGATGGAGGCTAA 240
DB	181 TTCAGTGGCAGTGGGTCTGGGACCTTACTCTCTCACAAATCAGCAGATGGAGGCTAA 240
OY	241 GATGTGTCACATTATTACTGCCAGCAGTAGGAGTACTAACCCACCCAGTTGGAGGGGG 300
DB	241 GATGTGTCACATTATTACTGCCAGCAGTAGGAGTATTAAACCAGCAGCTTGGTGGAGGC 300
OY	301 ACCATGCTGGAAATAAGA 318
DB	301 ACCAAGCTGGAAATCAAA 318
RESULT 11	
ABK24004	
ID	ABK24004 standard; cDNA; 335 BP.
XX	
AC	ABK24004:
XX	
DT	09-APR-2002 (first entry)
XX	
DE	Murine BC2 light chain modified variable region DNA.
XX	
KM	Human; mouse; BC2; animal post-thromboembolic induced ischaemia;
KW	thrombolytic agent; anti-factor IX antibody; plasminogen activator; ss;
KM	chromobolytic stroke; cerebroprotective; anticoagulant; Thrombolytic;
KW	vastotropic; cardiant; PCR primer; anti-respiratory syncytial virus;
XX	heavy chain variable region; light chain variable region.
OS	Mus sp.
OS	Synthetic.
PN	WO200187339-A1.
XX	
PD	22-NOV-2001.
XX	
PX	05-OCT-2000; 2000WO-US27438.
PR	15-MAY-2000; 2000US-0571434.
XX	
PA	(SMIK ) SMITHKLINE BEECHAM CORP.
PI	Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;
XX	
DR	WPI; 2002-082944/11.
DR	P-PsDB; AAU81001.
PT	Treating post-thromboembolic induced ischaemia in an animal by
PT	administering anti-factor IX antibody in combination with a plasminogen
XX	activator
PS	Example 7; Page 152-153; 163pp; English.



QY 301 ACCATGCTGGAATPAGA 318  
|||||  
Db 386 ACCAAGCTGAGCTGAAA 403

## RESULT 13

AAN91147  
ID AAN91147 standard; DNA; 403 BP.

AC AAN91147;

XX 06-JUL-1990 (first entry)

DE 2H7 Vh sequence.

XX Antibodies; passive immunisation; pH3-6a; ss.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 20..403

FT misc\_feature 320..352

FT /tag= a

FT /tag= b

FT /note="Jk5 region."

PN M08900999-A.

XX 09-FEB-1989.

PF 25-JUL-1988; 88WO-US02514.

PR 24-JUL-1987; 87US-0077528.

DR P-PSDB: AAP94778.

XX WPI: 1989-061144/08.

XX P-PSDB: AAP94778.

XX (ITGE-) INT GENETIC ENG INC.

XX Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;

XX WPI: 1989-061144/08.

XX P-PSDB: AAP94778.

XX (ITGE-) INT GENETIC ENG INC.

XX Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;

XX WPI: 1989-061144/08.

XX P-PSDB: AAP94778.

XX (ITGE-) INT GENETIC ENG INC.

XX Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;

XX WPI: 1989-061144/08.

XX P-PSDB: AAP94778.

XX (ITGE-) INT GENETIC ENG INC.

XX Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;

XX WPI: 1989-061144/08.

XX P-PSDB: AAP94778.

XX (ITGE-) INT GENETIC ENG INC.

Db 266 TTCACTGGCAGTGGCTGGAGCCTTACTCTCTACAAATCAGCAGAGTGAGGCTGAA 325  
QY 241 GATGCTGGCAGCTTATTACTGACAGTGGAGTAGTAAACCCAGCTGGAGGGG 300  
Db 326 GATGCTGGCAGCTTATTACTGACAGTGGAGTGGAGTTTAAACCCAGCTGGAGGCTGG 385  
QY 301 ACCATGCTGGAATPAGA 318  
|||||  
Db 386 ACCAAGCTGAGCTGAAA 403

## RESULT 14

AAV82358  
ID AAV82358 standard; DNA; 404 BP.

AC AAV82358;

XX 30-MAR-1999 (first entry)

DE Mouse antibody 2H7 light chain variable region encoding sequence.

XX pelB peptate lyase; secretion signal; chimeric antibody;

XX light chain; B-cell antigen; antibody 2H7; ss.

OS Mus sp.

XX Key Location/Qualifiers

FT CDS 21..404

FT /tag= a

FT /note="partial CDS"

PN US5846818-A.

XX 08-DEC-1998.

PF 06-JUN-1995; 95US-0472696.

PR 29-MAR-1990; 90US-0501092.

PR 01-NOV-1985; 85US-0793980.

PR 24-JUL-1987; 87US-0077528.

PR 11-JAN-1988; 88US-0142039.

PR 08-DEC-1992; 92US-0987555.

PR 22-FEB-1993; 93US-0020671.

PR 09-DEC-1994; 94US-0357234.

PR 06-JUN-1995; 95US-0472696.

XX (XOMA ) XOMA CORP.

XX Better M, Horowitz AH, Lei S, Liu AY, Robinson RR;

XX Wall R, Wilcox GL;

XX WPI: 1999-059072/05.

XX P-PSDB: AAM89541.

XX pelB peptate lyase signal sequence - and vector for expression of

XX secreted proteins in Gram-negative bacteria

XX Example 4; Fig 22; 98pp; English.

XX The present sequence encodes the mouse antibody 2H7 light chain

XX variable region. Antibody 2H7 is specific for human B-cell antigen. The

XX antibody sequence was used to construct a chimeric human-mouse antibody,

XX in the course of the invention. The chimeric antibody is expressed in a

XX secretion vector comprising a pelB peptate lyase secretion signal

XX peptide. The pelB peptate lyase secretion sequence is useful for

XX producing a protein such as a chimeric antibody in a bacterial host.

XX Sequence 404 BP; 100 A; 113 C; 93 G; 98 T; 0 other;

XX Query Match 93.0%; Score 295.6; DB 20; Length 404;

XX Best Local Similarity 95.6%; Pred. No. 8.7e-76;

XX Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGAGTCTCAGCAATCCTGTGCATCTCCAGGGGAAAGTTCACA 60  
 |||||  
 Db 87 CAAATGTTCTCTCCAGAGTCTCAGCAATCCTGTGCATCTCCAGGGGAAAGTTCACA 146  
 QY 61 ATGACTTGACAGGGGCGACCTCAAGTGTAAATTAATGACAGTGTACAGCAGCAAGCCAGCA 120  
 |||||  
 Db 147 ATGACTTGACAGGGGCGACCTCAAGTGTAAATTAATGACAGTGTACAGCAGCAAGCCAGCA 206  
 QY 121 TCCGCCCCCAAAACCTGGATTTCTGCCACATCCCAACCTGGTCTTGAGAGTCCCTGCTCGC 180  
 |||||  
 Db 207 TCCGCCCCCAAAACCTGGATTTATGCCCCCAACCTGGCTTGTGAGAGTCCCTGCTCGC 266  
 QY 181 TTCAGTGGCAGAGTGGTCTGGAGCCTTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 240  
 |||||  
 Db 267 TTCAGTGGCAGAGTGGTCTGGAGCCTTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 326  
 QY 241 GATGCTGCCACTTATTAATCTGCGCAGCAGTGGAGTAAACCCAGCCTTGGAGGGGG 300  
 |||||  
 Db 327 GATGCTGCCACTTATTAATCTGCGCAGCAGTGGAGTAAACCCAGCCTTGGAGGGGG 386  
 QY 301 ACCATGCTGGAATTAAGA 318  
 |||||  
 Db 387 ACCAAGCTGGAGCTGAAA 404

## RESULT 15

AAH22071 standard; DNA: 404 BP.

AAH22071:

17-AUG-2001 (first entry)

247 light chain variable region gene sequence.

Human; mouse; chimeric immunoglobulin; chimeric antibody;

genetic engineering; primer; ss.

Homo sapiens.

Mus sp.

US6204023-B1.

20-MAR-2001.

06-JUN-1995; 95US-0472691.

18-AUG-1994; 94US-0299085.

29-MAR-1990; 90US-0501092.

08-DEC-1992; 92US-0987555.

25-MAY-1995; 95US-0450731.

01-NOV-1985; 85US-0793980.

27-OCT-1986; 86MO-US02269.

24-JUL-1987; 87US-0077528.

(XOMA ) XOMA LTD.

Robinson RR, Liu AY, Horwitz AH, Better M, Wall R, Lei S;

Wilcox GL;

WPI; 2001-289514/30.

P-PSDB; AAB98093.

Preparing immunoglobulins or genetically engineered antibodies for the

large scale production of antibodies involves modular assembly of

antibodies through cloning and expression of light and heavy chains -

CC genetically engineered organisms. The method comprises: (a) expressing,  
 CC in a bacterial cell, of a first nucleic acid that encodes a bacterial  
 CC signal sequence operably linked to either the heavy chain or heavy chain  
 CC fragment, and a second nucleic acid that encodes a bacterial signal  
 CC sequence operably linked to either the light chain or light chain  
 CC fragment, and obtaining the immunoglobulin from the periplasmic space or  
 CC culture medium; or (b) operably linking a nucleic acid encoding a  
 CC bacterial signal sequence to a nucleic acid encoding the heavy chain  
 CC variable region or the light chain variable region or both of the  
 CC variable regions, and obtaining the immunoglobulin from the periplasmic  
 CC space or culture medium. The bacterial signal sequences bring about the  
 CC transport of the heavy chain or heavy chain fragment, and the light chain  
 CC or light chain fragment, or the variable regions through the cytoplasmic  
 CC membrane of a bacterial cell. The immunoglobulin comprises: (a) a heavy  
 CC chain or heavy chain variable region and a light chain or light chain fragment;  
 CC or (b) a heavy chain variable region and a light chain variable region.  
 CC The method is useful for producing genetically engineered antibodies of  
 CC desired variable region specificity and constant region properties. The  
 CC method is also useful large scale production of human antibodies.  
 CC AAH21985 to AAH22082 and AAB98085 to AAB98097 represent sequences used in  
 CC the exemplification of the present invention.

SQ Sequence 404 BP; 113 A; 113 C; 93 G; 98 T; 0 other;

Query Match 93.0%; Score 295.6; DB 22: Length 404;

Best Local Similarity 95.6%; Pred. No. 8.7e-76; Indels 0; Gaps 0;

Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGAGTCTCAGCAATCCTGTGCATCTCCAGGGGAAAGTTCACA 60  
 |||||  
 Db 87 CAAATGTTCTCTCCAGAGTCTCAGCAATCCTGTGCATCTCCAGGGGAAAGTTCACA 146  
 QY 61 ATGACTTGACAGGGGCGACCTCAAGTGTAAATTAATGACAGTGTACAGCAGCAAGCCAGCA 120  
 |||||  
 Db 147 ATGACTTGACAGGGGCGACCTCAAGTGTAAATTAATGACAGTGTACAGCAGCAAGCCAGCA 206  
 QY 121 TCCGCCCCCAAAACCTGGATTTCTGCCACATCCCAACCTGGTCTTGAGAGTCCCTGCTCGC 180  
 |||||  
 Db 207 TCCGCCCCCAAAACCTGGATTTATGCCCCCAACCTGGCTTGTGAGAGTCCCTGCTCGC 266  
 QY 181 TTCAGTGGCAGAGTGGTCTGGAGCCTTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 240  
 |||||  
 Db 267 TTCAGTGGCAGAGTGGTCTGGAGCCTTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 326  
 QY 241 GATGCTGCCACTTATTAATCTGCGCAGCAGTGGAGTAAACCCAGCCTTGGAGGGGG 300  
 |||||  
 Db 327 GATGCTGCCACTTATTAATCTGCGCAGCAGTGGAGTAAACCCAGCCTTGGAGGGGG 386  
 QY 301 ACCATGCTGGAATTAAGA 318  
 |||||  
 Db 387 ACCAAGCTGGAGCTGAAA 404

Search completed: November 27, 2002, 03:53:51  
 Job time : 187.766 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:45:51 ; Search time 1318.29 Seconds  
(without alignments)  
3906.706 Million cell updates/sec

Title: US-09-893-615-88

Sequence: 1 CAAATGTCTCTCCAGTC.....GGACCACTCGAATAAGA 318

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estnu:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	286.8	90.2	639	13	BG968615 602836416
2	276.4	86.9	959	14	BQ939046 AGENCOURT
3	274.8	86.4	532	13	B1104783 B1104783
4	274.4	86.3	798	13	BG968518 602835104
5	273.2	85.9	755	13	B1150509 602915167
6	273.2	85.9	918	14	BQ946795 AGENCOURT

7	272.2	85.6	312	12	BG090240 uc58e10.y
8	271.2	85.3	413	10	BE624609
9	267.2	84.0	376	9	AA691311 vs14f01.t
10	266.8	83.9	891	12	BF165585
11	266.6	83.8	865	13	BG964349
12	266.4	83.8	960	12	BE581607
13	263.6	82.9	728	13	BG962985
14	263.2	82.8	639	12	BF583310
15	262.2	82.5	743	13	B1106381
16	260.4	81.9	472	10	BE289692
17	260.4	81.9	862	13	BG962582
18	259.4	81.6	975	13	BG961850
19	257.8	81.1	948	12	BF139176
20	256.2	80.6	876	12	BF141750
21	256	80.5	807	17	BH035104
22	253	79.6	1532	12	BF133688
23	251.4	79.1	918	12	BF142385
24	250.8	78.9	641	12	BF134573
25	246.8	77.6	569	13	BG964957
26	246.6	77.5	401	12	BF015569
27	246.6	77.5	854	14	BQ947692
28	245.8	77.3	722	12	BF140170
29	245.8	77.3	1000	14	BQ937284
30	245.6	77.2	953	12	BE583869
31	245	77.0	403	14	BQ375976
32	245	77.0	407	14	BQ328151
33	245	77.0	409	12	BF822649
34	245	77.0	410	12	BF768589
35	245	77.0	416	13	B1025870
36	245	77.0	419	14	BQ321312
37	245	77.0	424	14	BQ376239
38	245	77.0	425	13	B1063029
39	245	77.0	427	12	BF753315
40	245	77.0	432	12	BF843458
41	245	77.0	433	12	BF843458
42	245	77.0	437	14	BQ321307
43	245	77.0	439	14	BQ321297
44	245	77.0	444	12	BF894856
45	245	77.0	444	12	BF894857

#### ALIGNMENTS

RESULT 1  
LOCUS BG968615 639 bp mRNA linear EST 12-JUN-2001  
DEFINITION 602836416F1 NCI\_CGAP\_C024 Mus musculus cDNA clone IMAGE:4991067 5',  
mRNA sequence.

ACCESSION BG968615  
VERSION BG968615.1 GI:14356252

KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 639)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-riemail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M11007 row: O column: 04  
High quality sequence stop: 639.  
Location/Qualifiers  
1. 639

#### FEATURES

source

/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4991067"  
/clone\_lib="NCI CGAP Co24"  
/lab\_host="DH10B (71 phage-resistant)"  
/note="Organ: colon; Vector: pcwv-SPOrt6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 172 a 173 c 151 g 143 t  
ORIGIN

Query Match 90.2%; Score 286.8; DB 13; Length 639;  
Best Local Similarity 95.9%; Pred. No. 2.5e-75;  
Matches 305; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 CAATTTCTTCTCCAGTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGGTACA 60  
|||||  
Db 76 CAATTTCTTCTCCAGTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGGTACA 134  
QY 61 ATGACTTGCAGGGCCAGCTCAAGTGAATTAATACATGCACTGTACAGCAGAAAGCCAGGA 120  
|||||  
Db 135 ATGACTTGCAGGGCCAGCTCAAGTGAATTAATACATGCACTGTACAGCAGAAAGCCAGGA 194  
QY 121 TCTCTCCCCAAACCTGGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180  
|||||  
Db 195 TCTCTCCCCAAACCTGGATTTATGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 254  
QY 181 TTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAGATCAGCAGAGTGGAGGCTGAA 240  
|||||  
Db 255 TTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAGATCAGCAGAGTGGAGGCTGAA 314  
QY 241 GATGCTGCCACTTATTAATCTCCAGCACTGGAGTGAATACCCAGCTTCGGAGGGGG 300  
|||||  
Db 315 GATGCTGCCACTTATTAATCTCCAGCACTGGAGTGAATACCCAGCTTCGGAGGGGG 374  
QY 301 ACCATGCTGGAATTAAGA 318  
|||||  
Db 375 ACCAAGCTGGAATTAAGA 392

RESULT 2 959 bp mRNA linear EST 21-AUG-2002  
B0939046  
LOCUS  
DEFINITION  
AGENCOURT 8946838 NCI CGAP Co24 Mus musculus cDNA clone  
IMAGE:6395050 5', mRNA sequence.

ACCESSION B0939046  
VERSION B0939046.1 GI:22354524  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 959)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LHAM13889 row: b column: 11  
High quality sequence stop: 517.

## FEATURES

source  
1..959  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="FVB/N"

/db\_xref="taxon:10090"  
/clone="IMAGE:6395050"  
/clone\_lib="NCI CGAP Co24"  
/lab\_host="DH10B (71 phage-resistant)"  
/note="Organ: colon; Vector: pcwv-SPOrt6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 267 a 256 c 214 g 222 t  
ORIGIN

Query Match 86.9%; Score 276.4; DB 14; Length 959;  
Best Local Similarity 91.8%; Pred. No. 4.2e-72;  
Matches 292; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 CAATTTCTTCTCCAGTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGGTACA 60  
|||||  
Db 70 CAATTTCTTCTCCAGTCTCCAGCAATCTCTGTCATCTCCAGGGGAAAGGTACC 129  
QY 61 ATGACTTGCAGGGCCAGCTCAAGTGAATTAATACATGCACTGTACAGCAGAAAGCCAGGA 120  
|||||  
Db 130 ATGACTTGCAGGGCCAGCTCAAGTGAATTAATACATGCACTGTACAGCAGAAAGCCAGGA 189  
QY 121 TCTCTCCCCAAACCTGGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180  
|||||  
Db 190 ACCTCCCCAAAGAGTGAATTAATGACATCCAGACTGGCTTCTGGAGTCCCTGCTCGC 249  
QY 181 TTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAGATCAGCAGAGTGGAGGCTGAA 240  
|||||  
Db 250 TTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAGATCAGCAGAGTGGAGGCTGAA 309  
QY 241 GATGCTGCCACTTATTAATCTCCAGCACTGGAGTGAATACCCAGCTTCGGAGGGGG 300  
|||||  
Db 310 GATGCTGCCACTTATTAATCTCCAGCACTGGAGTGAATACCCAGCTTCGGAGGGGG 369  
QY 301 ACCATGCTGGAATTAAGA 318  
|||||  
Db 370 ACCAAGCTGGAATTAAGA 387

RESULT 3 532 bp mRNA linear EST 26-JUN-2001  
B1104783  
LOCUS  
DEFINITION  
602891329F1 NCI CGAP Lu29 Mus musculus cDNA clone IMAGE:5036622 5',  
mRNA sequence.

ACCESSION B1104783  
VERSION B1104783.1 GI:14555676  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 532)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LHAM11101 row: 1 column: 07  
High quality sequence stop: 532.

## FEATURES

source  
1..532  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5036622"



/clone.lib="NCI\_CGAP\_Lu29"  
/tissue\_type="spontaneous tumor, metastatic to mammary."  
Stem cell origin  
/lab\_host="DH10B"  
/note="Organ: Lung; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

BASE COUNT 139 a 148 c 113 g 132 t  
ORIGIN

Query Match 86.4%; Score 274.8; DB 13; Length 532;  
Best Local Similarity 91.5%; Pred. No. 9.3e-72;  
Matches 291; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGTCTCCAGCAATCTGTCTGCATCTCCAGGGGAAAAGTCA 60  
DB 102 CAAATTTCTCTCAGCTCAGCTCAGCAATCATGTCTGCATCTCCAGAGAGAGTCA 161  
QY 61 ATGACTTGAGGGGCGAGTCAAGTGAATTAATGACACTGTGACAGAGAGAGAGAG 120  
DB 162 ATGACCTGAGTGCAGCTCAGTGAATTAATGACACTGTGACAGAGAGAGAGAG 221  
QY 121 TCTCTCCCAAAACCTGGATTCTGCAATCCAACTGGCTTGGAGTCCCTGCTGC 180  
DB 222 TCTCTCCCAAAACCTGGATTCTGCAATCCAACTGGCTTGGAGTCCCTGCTGC 281  
QY 181 TTCACTGAGAGTGGTCTGGAGCTCTTACTCTCTCAATCAGAGAGAGAGAGTGA 240  
DB 282 TTCACTGAGAGTGGTCTGGAGCTCTTACTCTCTCAATCAGAGAGAGAGAGTGA 341  
QY 241 GATGTGCACTTATTACTGCGAGAGTGAATCAACCCAGCTGGAGGGGGG 300  
DB 342 GATGTGCACTTATTACTGCGAGAGTGAATCAATGTTACCCGTAATGTTGGAGTGG 401  
QY 301 ACCATGCTGGAATAGA 318  
DB 402 ACCAAGCTGGAATAGAA 419

RESULT 4  
BC968518 798 bp mRNA linear EST 12-JUN-2001  
LOCUS 602835104F1 NCI\_CGAP\_CO24 Mus musculus cDNA clone IMAGE:4989720 5',  
DEFINITION mRNA sequence.  
ACCESSION BC968518  
VERSION BC968518.1 GI:14356142  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 798)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM11004 row: g column: 01  
High quality sequence stop: 797.  
Location/Qualifiers  
1. 798  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAG:4989720"

FEATURES  
source

/clone.lib="NCI\_CGAP\_CO24"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP library."

BASE COUNT 225 a 222 c 179 g 172 t  
ORIGIN

Query Match 86.3%; Score 274.4; DB 13; Length 798;  
Best Local Similarity 91.8%; Pred. No. 1.5e-71;  
Matches 290; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 3 AATTGTTCTCTCCAGTCTCCAGCAATCTGTCTGCATCTCCAGGGGAAAAGTCA 62  
DB 86 AATGTTCTCACCAGCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAGTCA 145  
QY 63 GACTTGACAGGGGCGAGTCAAGTGAATTAATGACACTGTGACAGAGAGAGAGATC 122  
DB 146 GACTTGACAGGGGCGAGTCAAGTGAATTAATGACACTGTGACAGAGAGAGAGATC 205  
QY 123 CTCCCAAAACCTGGATTCTGCAATCCAACTGGCTTGGAGTCCCTGCTGC 182  
DB 206 CTCCCAAAACCTGGATTCTGCAATCCAACTGGCTTGGAGTCCCTGCTGC 265  
QY 183 CAGTGGCACTGGGTCTGGAGCTCTTACTCTCTCAATCAGAGAGAGAGAGTGA 242  
DB 266 CAGTGGCACTGGGTCTGGAGCTCTTACTCTCTCAATCAGAGAGAGAGAGTGA 325  
QY 243 TCTGCGCACTTATTACTGCGAGAGTGAATCAACCCAGCTGGAGGGGGG 302  
DB 326 TCTGCGCACTTATTACTGCGAGAGTGAATCAATGTTACCCGTAATGTTGGAGTGG 385  
QY 303 CATGCTGGAATAGA 318  
DB 386 CAAGCTGGAATAGAA 401

RESULT 5  
B1150509 755 bp mRNA linear EST 05-JUL-2001  
LOCUS 602915167F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:5065720 5',  
DEFINITION mRNA sequence.  
ACCESSION B1150509  
VERSION B1150509.1 GI:14610510  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 755)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM11177 row: e column: 17  
High quality sequence stop: 738.  
Location/Qualifiers  
1. 755  
/organism="Mus musculus"  
/strain="CZECH IT"  
/db\_xref="taxon:10090"  
/clone="IMAG:5065720"  
/clone.lib="NCI\_CGAP\_Lu29"  
/tissue\_type="spontaneous tumor, metastatic to mammary."

FEATURES  
source

```

Stem cell origin."
/lab_host="DH10B"
/node="Organ: Lung; Vector: pCMV-SPORT6; Site_1: SalI,
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT
library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

Query Match	Similarity	85.9%	Score 273.2	DB 13	Length 755
Best Local	Similarity	91.2%	Pred. No. 3.4e-71		
Matches	290	Conservative	0	Mismatches	28
				Indels	0
				Gaps	0
Qy	1	CAAAATGTTCTCTCCAGCTCCACAAATCCCTGTCTGCATCTCCAGGGGAAAAGSTCACA	60		
Db	84	CAAAATGTTCTCCACCAAGTCTCAGACATTCATGTTGATCTCCGGGGGAGAAATCACC	143		
Qy	61	ATGACTTGACAGGGCCAGCTCAAGTGAATTATACATGACACTGTGTACACAGACCAAGA	120		
Db	144	ATMACTGTGACAGTCCAGCAGCTAAGTTACATGACACTGGTATACAGACAGAACCAAGA	203		
Qy	121	TCCCTCCCCAAACCCGTGATTTCTGCCAATCCAAACGCGCTTCGAGATCCCTCTCGC	180		
Db	204	TCTCTCCCCAAACCTGTGGATTTATACACATCCAAACTGGCTCTCGAGATCCCTCTCGC	263		
Qy	181	TTCAGTGGCAGAGGGTCTGGGACCTCTTACTCTCACCATGACAGAGTGGAGCTGAA	240		
Db	264	TTCAGTGGCAGAGGGTCTGGGACCTCTTACTCTCACCATGACAGCATGAGAGCTGAA	323		
Qy	241	GATGCTGCCACTTATTACTGCCAGAGTGGAGTACTAAACCCACCCAGCTTGGAGGGGGG	300		
Db	324	GATGCTGCCCTTATTATTCTGGCCATCAGTGGAGTACTTACCACCAACGTTCCGATCGGGG	383		
Qy	301	ACCATGCTGGAAATAAGA	318		
Db	384	ACCATGCTGGAAATAAAA	401		

[illegible]

## FEATURES

### source

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="IMAGE:6474845"
/clone_lib="NCI-CGAP Co24"
/lab_host="DH10B (T1 phage-resistant)"

```

BASE COUNT	246 a	257 c	193 g	212 t	10 others
ORIGIN					
Query Match	85.9%	Score 273.2	DB 14	Length 918	
Best Local Similarity	91.2%	Pred. No. 3,7e-71			
Matches 290	Conservative	0	Mismatches 20	Indels 0	Gaps 0

OY	1	CAAAATGTTCTCTCCAGCTCCAGCTCCAGCAATCCCTGTGTGATCTTCACAGGGGCAAAAGGTCACA	60
Db	82	CAAAATGTTCTCTCCAGCTCCAGCTCCAGCAATCATGTCTGCATCTCCAGGGGAGAAAGTCAAC	141
OY	61	ATGACTTTCAGGGCCAGCTCAAGGTAAATTTCATATGCACTGGTACCACAGAAAGCGCAGA	120
Db	142	ATGACCTTCAGAGGCCAGCTCAAGTATAGTTCTCAATGCACTGGTATCCACGAGAAATGACGC	201
OY	121	TCTCTCCGCCCAAAACCTGGATTTCTTGTGCAATCCAACTGGCTTTCGAGATCCCTGCTGCG	180
Db	202	ACCTCCGCCCAAAAGATGGATTTATGACACATCCAAATATGGCTTTCGAGATCCCTGCTGCG	261
OY	181	TTCACTGTCAGTGGGTCTGCGACCTTCTACTCTTCACAAATAGCAGATGGAGGCTGAA	240
Db	262	TTCACTGTCAGAGGGGTCTGCGACCTTCTACTCTTCACAAATAGCAGATGGAGGCTGAA	321
OY	241	GATGCTGCACACTATTACTGCGCCAGCAGTAGTGAGTAACCCACCCACAGTTCGAGAGGGGG	300
Db	322	GATGCTGCACACTATTACTGCGCCAGCAGTAGTGAGTAAGAACCCACCCACAGTTCGCGTCTGGGG	381
OY	301	ACCATGTCGGAATTAAGA	318
Db	382	ACAAAGTTGGAAATATAAA	399

RESULT 7  
 BG090240  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 FEATURES  
 source

BG090240 312 bp mRNA linear EST 26-JAN-2001  
 uc58be10.y1 Soares\_mouse\_MNGB\_bcell Mus musculus cDNA clone  
 IMAGE:3333082 5' similar to TF:G9U410 G9U410 MONOCLONAL  
 ANTI-IDIOCYPIC ANTIBODY NP30 IMMUNOGLOBULIN LIGHT CHAIN VARIABLE  
 REGION ;, mRNA sequence.  
 BG090240  
 BG090240.1 GI:12572803  
 EST.  
 house mouse.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 312)  
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CCGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 MGI:1076246  
 Seq primer: -40RP from Gibco.  
 Location/Qualifiers  
 1..312

**FEATURES**  
**source**

```
/organism="Mmus musculus"  
/db_xref="taxon:10090"  
/clone_image=3332082"  
/_clone_1lb="Soares_mouse_NK6B_bcell)"  
/_lab_host="DH10B (phage-resistant)"  
/_note="Organ: geminal B-cell; Vector: pT7AD-Pac  
(pharmacia) with a modified polylinker; Site I:  
Not I ; Site-2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5'  
tgtttaccactctcaatgtggagcgccgacctgcattt]";
```



```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image:118201"
/clone_id="Barstead mouse irradiated colon MPluR7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoRI. Site_2: NotI. Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - o1igo(dT) primer
[5'TGTACGATCTGAAAGTGAGCGCCGCCCTTTTTTTTTTTTTTTTTTTT
T 3']: double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGATCCGTC], digested with NotI and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library constructed by Bob Barstead."

```

Query Match	Similarity	84.0%;	Score 267.2;	DB 9;	Length 396;
Best Local	Similarity	92.4%;	Pred. No. 1.5e-69;		
Matches	281;	Conservative	0;	Mismatches 23;	Indels 0;
					Gaps 0;
OY	1	CAAAATGTTCTCTCCAGCTCTCCAGCAATCTCTGTCTGCATCTCCAGGGGAAAAGTCA	60		
Db	92	CAAAATGTTCTCTCCAGCTCTCCAGCAATCTCTGTCTGCATCTCCAGGGGAAAAGTCA	151		
OY	61	ATGACTTTCAGGGCCAGCTCAAGTAAATTAATCATGCACTGGTACCACAGAACGACGA	120		
Db	152	ATGACTTTCAGGGCCAGCTCAAGTAAATTAATCATGCACTGGTACCACAGAACGACGA	211		
OY	121	TCCTCCCCCAAAACCCTGGAGTTCTGCGACATCCAACTGGCTTCGAGATCCCTCTGC	180		
Db	212	ACCTCCCCCAAAAGATGATTTATATACACATCCAAACGGCTTCGAGATCCCTCTGC	271		
OY	181	TTCACTGCGAGTGGGTCTGGGACCTTACTCTCTCACAAATCAGCAGATGGAGGCTGAA	240		
Db	272	TTCACTGCGAGTGGGTCTGGGACCTTACTCTCTCACAAATCAGCAGATGGAGGCTGAA	331		
OY	241	GATGCTGCGACTTATTACTGCGACGAGTGGAGTAACTCAACCCAGCTGGAGAGGGGG	300		
Db	332	GATGCTGCGACTTATTACTGCGACGAGTGGAGTAACTCAACCCAGCTGGAGAGGGGG	391		
OY	301	ACCA 304			
Db	392	ACCA 395			
RESULT 10					
LOCUS	BF165585	871 bp	mRNA	linear	EST 30-OCT-2000
DEFINITION	60177474861	NCI_GCAP_Lu29	Mus musculus	cdna clone	IMAGE:4019109 5',
ACCSSION	BF165585				
VERSION	BF165585.1	GI:11045937			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 871)				
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a>				
	Tissue Procurement: Gilbert Smith, Ph.D.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LMN at:				

BASE COUNT	ORIGIN	FEATURES
226	a	<p>http://image.lim1.gov</p> <p>Plate: ILM09270 row: 1 column: 22</p> <p>High quality sequence stop: 714.</p> <p>Location/Qualifiers</p> <p>1. .871</p> <p>/organism="Mus musculus"</p> <p>/strain="CZECH 11"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:4019109"</p> <p>/clone_1ib="NCI CGAP Lu29"</p> <p>/tissue_type="Spontaneous tumor, metastatic to mammary.</p> <p>Stem cell origin"</p> <p>/lab_host="DH10B"</p> <p>/note="Organ: lung; Vector: PCMV-SPOPT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. library constructed by life Technologies. Investigator providing samples: Gilbert Smith, NIH"</p>
226	c	201 g 204 t

Query Match	83.9%	Score 266.8	DB 12	Length 871
Best Local Similarity	89.9%	Pred. No. 3e-69		
Matches 286	Conservative 0	Mismatches 32	Indels 0	Gaps 0
QY	1	CAAAATGTTCTCTCCAGCTCTCCAGCAATCTGTCTGATCTCCAGGGAAAGTCA	60	
DB	90	CAAAATGTTCTCTCCAGCTCTCCAGCAATCAATGATCTGATCTCCAGGGAAAGTCA	149	
QY	61	ATGACTTGAGGGCCAGCTCAGTGAATTAATGACACTGGTACCAGCAAGCCAGGA	120	
DB	150	ATGACCTGAGAGTCCAGCTCAAGTGAATTAATGACACTGGTATCAGCAAGCCAGGA	209	
QY	121	TGCTTCCCCCAACCCCTGGATTTCTGGCACATCCAACTGGCTCTTGAGATCCCTGCTCG	180	
DB	210	TGCTTCCCCCAACCCCTGGATTTATGTTACATCCGACCTGGCTCTTGAGATCCCTGCTCG	269	
QY	181	TTTCAGTGGCAGTGGGCTCTGGACCTTTACTCTCTCAATACAGCAAGATGGAGCTGAA	240	
DB	270	TTTCAGTGGGAGTGGGAGTGGGACCTTTACTCTCTCAATACAGCAAGATGGAGCTGAA	329	
QY	241	GATGTCGCCACTTATTAACGACGACGAGTGGAGTGAAGTGAACCCACGCTTGGAGGGGG	300	
DB	330	GATGTCGCCACTTATTTCTGCCAGCAGTATCATATGTTACCCACGACGCTTGGTGGAGGC	389	
QY	301	ACCATGCTGGAAATAGA	318	
DB	390	ACCAAGCTGGAAATAGAA	407	
RESULT 11				
LOCUS	BG964349	865 bp	mRNA	linear
DEFINITION	602831966E1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4986568 5'			
ACCESSION	BG964349			
VERSION	BG964349.1			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (bases 1 to 865)			
TITLE	NIH-MGC http://mgc.nci.nih.gov/			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgab@remail.nih.gov			
	Tissue Procurement: Jeffrey E. Green, M.D.			
	cDNA Library Preparation: Life Technologies, Inc.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:			



## FEATURES

Location/Qualifiers  
1. 728/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4983018"/clone\_1lib="NCI\_CGAP\_Co24"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 202 a 198 c 168 g 160 t

## ORIGIN

Query Match 82.9%; Score 263.6; DB 13; Length 728;

Best Local Similarity 89.3%; Pred. No. 2.5e-68;  
Matches 284; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1 CAATATTGTTCTCTCCAGTCTCCAGCAATCTCTGCATCTCCAGGGGAAAGGTCACA 60  
|||||  
Db 54 CAATATTGTTCTCACACGAGTCTCCAGCAATCATCTCTGCATCTCCAGGGGAAAGGTCACC 113  
|||||  
QY 61 ATACATTGCGAGGGGCGAGCTCAATGTAATATACATGACACGACGAGAAAGCCAGGA 120  
|||||  
Db 114 ATGACCTGCGAGGCGAGCTCAATGTAATATACATGACGAGGAGGAGGAGG 173  
|||||  
QY 121 TCTCCGCCCAAAACCTGGATTTCTGCCACATCCAGCTGCTTGGAGTCCCTGCTCGC 180  
|||||  
Db 174 ACCTCCGCCCAAAAGATGATTTATGACACATCCAACTGCTTCTGAGTCCCTGCTCGC 233  
|||||  
QY 181 TTCACTGGGCGAGTGGTGGGACCTCTTACTCTCTCAATCAATCAAGAGAGGCTGAA 240  
|||||  
Db 234 TTCACTGGGCGAGTGGTGGGACCTCTTACTCTCTCAATCAATCAAGAGAGGCTGAA 293  
|||||  
QY 241 GATGCTGCGCACTTATATCTCCAGCAATGTAATGTAATGTAATGTAATGTAATGTAAT 300  
|||||  
Db 294 GATGCTGCGCACTTATATGTCATGAGCGAGGAGTATGAGGAGGAGGAGGAGG 353  
|||||  
QY 301 ACCATGCTGGAATTAAGA 318  
|||||  
Db 354 ACCATGCTGGAATTAAGA 371  
|||||

RESULT 14 639 bp mRNA linear EST 12-DEC-2000  
BF583310  
LOCUS 602101781F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4224692 5',  
DEFINITION mRNA sequence.  
ACCESSION BF583310  
VERSION BF583310  
KEYWORDS GI:11657028  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL 1 (bases 1 to 639)  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM9815 row: b column: 21  
High quality sequence stop: 636.

FEATURES  
source Location/Qualifiers  
1. 639/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4224692"/clone\_1lib="NCI\_CGAP\_Co24"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 169 a 174 c 149 g 147 t

## ORIGIN

Query Match 82.8%; Score 263.2; DB 12; Length 639;

Best Local Similarity 89.6%; Pred. No. 3.1e-68;  
Matches 283; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 3 AATGTTCTCTCCAGTCTCCAGCAATCTCTGCATCTCCAGGGGAAAGGTCACAAT 62  
|||||  
Db 78 AATGTTCTCTCACACGAGTCTCCAGCAATCATCTCTGCATCTCCAGGGGAAAGGTCACC 137  
|||||  
QY 63 GACTTCAGGCGGAGCTCAAGTGAATATACATGACATGCTGACACAGCAGGAGATC 122  
|||||  
Db 138 ATCTGCGAGTGGCGAGCTCAAGTGAATATACATGACATGCTGACACAGCAGGAGATC 197  
|||||  
QY 123 CTCCGCCCAAAACCTGGATTTCTGCCACATCCAACTGCTTGGAGTCCCTGCTCGCT 182  
|||||  
Db 198 CTCCGCCCAAAACCTGGATTTATGCAATCCAGCTGCTTGGAGTCCCTGCTCGCT 257  
|||||  
QY 183 CAGTGGCGAGTGGTGGGACCTTACTCTCTCAATCAATGAGAGGAGGCTGAGA 242  
|||||  
Db 258 CAGTGGCGAGTGGTGGGACCTTACTCTCTCAATCAATGAGAGGCTGAGA 317  
|||||  
QY 243 TCTGCGCACTTATATCTCCAGCAATGTAATGTAATGTAATGTAATGTAATGTAAT 302  
|||||  
Db 318 TCTGCGCACTTATATCTCCAGCAATGTAATGTAATGTAATGTAATGTAATGTAAT 377  
|||||  
QY 303 CATGCTGGAATTAAGA 318  
|||||  
Db 378 CAAGCTGAGCTGAAA 393  
|||||

RESULT 15 743 bp mRNA linear EST 26-JUN-2001  
B1106381  
LOCUS 602892928F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:5038110 5',  
DEFINITION mRNA sequence.  
ACCESSION B1106381  
VERSION B1106381  
KEYWORDS GI:14557274  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL 1 (bases 1 to 743)  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM1105 row: g column: 07  
High quality sequence stop: 728.

FEATURES  
source Location/Qualifiers  
1. 743  
/organism="Mus musculus"  
/strain="CZECH II"

/db\_xref="taxon:10090"  
/clone\_lib="NCI\_CGAP\_Lu29"  
/tissue\_type="spontaneous tumor, metastatic to mammary.  
stem cell origin"  
/lab\_host="DH10B"  
/note="Organ: Lung; Vector: pCMV-Sport6; Site:1; Salt:  
Site:2; Noted: Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"  
BASE COUNT 197 a 214 c 168 g 164 t  
ORIGIN

Query Match 82.5%; Score 262.2; DB 13; Length 743;  
Best Local Similarity 89.5%; Pred. No. 6.7e-68;  
Matches 282; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
OY 4 ATTGTTCTCTCCAGTCTCCAGCAATCCGTGTGCATCTCAGGGAAAGGTCAACATG 63  
DB 25 ATTCTTCACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGAGAGAGGTCAACATG 84  
OY 64 ACTTCAGGGCCAGCTCAAGTGAATTAATGATGACACTGGTACCAAGAGCCAGGATCC 123  
DB 85 ACCTGCAGTCCAGCTCAAGTGAATTAATTCATGCACTGGTATCAGCAGAGAGCCAGATCC 144  
OY 124 TCCCCAAACCTGGATTTCTGCCACATCCACACTGGCTTCTGGAGTCCCTGCTGCTTC 183  
DB 145 TCCCCAAAGCCCTGGATTTCTGTACATCCGACCTGGCTTCTGGAGTCCCTGCTGCTTC 204  
OY 184 AGTGCAGTGGGTCTGGAGCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGAT 243  
DB 205 AGTGCAGTGGAGCTGGAGCTCTTACTCTCTCACAATCAGCAGATGAGAGGCTGAAGAT 264  
OY 244 GCTGCCACTTATTACTGCCAGAGTGGAGTAGTAACCCACCGTTCGGAGGGGGGACC 303  
DB 265 GCTGCCACTTATTCTGCCAGAGTATCATAGTACCCAGGAGCGTTCGGTGGAGGACCC 324  
OY 304 ATGCTGAAATAGA 318  
DB 325 AAGCTGAAATCAAA 339

Search completed: November 27, 2002, 05:29:45  
Job time : 1336.29 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:35:53 ; Search time 35.6419 Seconds  
(Without alignments)  
2736.194 Million cell updates/sec

Title: US-09-893-615-88

Perfect score: 318

Sequence: 1 CAATGTCTCTCCAGTC.....GGACCATGCGAATAGA 318

Scoring table: IDENTITY\_NUC

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued\_patents.NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCrUS-COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300.4	94.5	321	US-08-783-853A-6	Sequence 6, Appl1
2	300.4	94.5	321	US-09-344-050-6	Sequence 6, Appl1
3	300.4	94.5	384	US-08-149-099C-6	Sequence 6, Appl1
4	300.4	94.5	384	US-08-478-967A-6	Sequence 6, Appl1
5	300.4	94.5	384	US-08-475-815B-6	Sequence 6, Appl1
6	300.4	94.5	9209	US-08-149-099C-3	Sequence 3, Appl1
7	300.4	94.5	9209	US-08-478-967A-3	Sequence 3, Appl1
8	300.4	94.5	9209	US-08-475-815B-3	Sequence 3, Appl1
9	300.4	94.5	18986	US-08-819-866-2	Sequence 2, Appl1
10	300.4	94.5	18986	US-09-023-715-2	Sequence 2, Appl1
11	300.4	94.5	18986	US-09-343-485A-2	Sequence 2, Appl1
12	300.4	94.5	384	US-08-476-275-3	Sequence 3, Appl1
13	297.2	93.5	9209	US-08-476-275-2	Sequence 2, Appl1
14	295.6	93.0	318	US-08-783-853A-104	Sequence 104, App
15	295.6	93.0	318	US-09-344-050-104	Sequence 104, App
16	295.6	93.0	335	US-08-783-853A-102	Sequence 102, App
17	295.6	93.0	335	US-09-344-050-102	Sequence 102, App
18	292.4	91.9	387	US-08-449-287-1	Sequence 1, Appl1
19	290.8	91.4	732	US-08-860-882A-26	Sequence 26, Appl1
20	290.8	91.4	732	US-09-423-439-57	Sequence 57, Appl1
21	290.8	91.4	732	US-09-011-769A-22	Sequence 22, Appl1
22	287.6	90.4	705	US-09-423-439-17	Sequence 17, Appl1
23	287.6	90.4	3217	US-09-423-439-52	Sequence 52, Appl1
24	285.8	89.9	315	US-08-459-310-3	Sequence 3, Appl1
25	281.8	88.6	321	US-08-211-202-4	Sequence 4, Appl1
26	281.8	88.6	711	US-08-061-092A-1	Sequence 1, Appl1
27	281.2	88.4	758	US-08-279-772A-5	Sequence 5, Appl1

28	281.2	88.4	759	US-08-902-486-8	Sequence 8, Appl1
29	280.2	88.1	711	US-08-468-252-4	Sequence 4, Appl1
30	280.2	88.1	711	US-08-668-706B-4	Sequence 4, Appl1
31	280.2	88.1	711	PCT-US95-10740-4	Sequence 4, Appl1
32	273.4	86.0	292	US-08-308-494A-12	Sequence 12, Appl1
33	273.2	85.9	384	US-08-656-586-1	Sequence 1, Appl1
34	273.2	85.9	393	US-08-116-778E-39	Sequence 39, Appl1
35	273.2	85.9	393	US-08-438-562-39	Sequence 39, Appl1
36	273.2	85.9	393	US-08-483-528B-2	Sequence 2, Appl1
37	273.2	85.9	393	US-08-673-799C-2	Sequence 2, Appl1
38	273.2	85.9	393	US-09-393-385B-2	Sequence 2, Appl1
39	271.8	85.5	292	US-09-280-028-3	Sequence 3, Appl1
40	271.4	85.3	331	US-08-836-561-32	Sequence 32, Appl1
41	270	84.9	669	US-08-190-199A-66	Sequence 66, Appl1
42	270	84.9	708	US-08-190-199A-60	Sequence 60, Appl1
43	268.4	84.4	390	US-08-482-882-79	Sequence 79, Appl1
44	268.4	84.4	390	US-08-483-389-79	Sequence 79, Appl1
45	268.4	84.4	390	US-08-487-113D-79	Sequence 79, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-783-853A-6  
Sequence 6, Application US/08783853A  
Patent No. 6005091  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Gloria  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
TITLE OF INVENTION: OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedekline Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/783,853A  
FILING DATE: 16-JAN-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/029,119  
FILING DATE: 24-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA

HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-783-853A-6

Query Match 94.5%; Score 300.4; DB 3; Length 321;  
Best Local Similarity 96.5%; Pred. No. 9e-86;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGCTCTCCAGCAATCTCTCTGCATCTCCAGGGGAAAAGTCCACA 60  
DB 1 CAAATGTTCTCTCCAGCTCTCCAGCAATCTCTCTGCATCTCCAGGGGAAAAGTCCACA 60  
QY 61 ATGACTTGACAGGGGCGAGCTCAAGTAAATTAATCATGACAGGGTACAGCAGAGCCAGGA 120  
DB 61 ATGACTTGACAGGGGCGAGCTCAAGTAAATTAATCATGACAGGGTACAGCAGAGCCAGGA 120  
QY 121 TCCTCCCGCCAAACCCGTGATTTCTGCGACATCCACCTGGCTTCTGGAGTCCCTGCTCGC 180  
DB 121 TCCTCCCGCCAAACCCGTGATTTCTGCGACATCCACCTGGCTTCTGGAGTCCCTGCTCGC 180  
QY 181 TTCAGTGGCAGTGGGTCTGGAGACCTTCTCTCTCATCATCAGCAGAGTGGAGCTGAA 240  
DB 181 TTCAGTGGCAGTGGGTCTGGAGACCTTCTCTCTCATCATCAGCAGAGTGGAGCTGAA 240  
QY 241 GATGCTGCGCATTATTACTGCGCAGCAGTGGAGTAAACCCAGCCAGTTCGAGGGGG 300  
DB 241 GATGCTGCGCATTATTACTGCGCAGCAGTGGAGTAAACCCAGCCAGTTCGAGGGGG 300  
QY 301 ACCAGTCTGGAAATTAAGA 318  
DB 301 ACCAGTCTGGAAATTAAGA 318

RESULT 2  
US-09-344-050-6  
Sequence 6, Application US/09344050  
Patent No. 6391299

GENERAL INFORMATION:

APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Gloria  
APPLICANT: Nichols, Andrew  
APPLICANT: Padian, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
TITLE OF INVENTION: OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/344, 050

FILING DATE: 24-JUN-1999

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/783,853

FILING DATE: 16-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Baumelster, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-09-344-050-6

Query Match 94.5%; Score 300.4; DB 4; Length 321;  
Best Local Similarity 96.5%; Pred. No. 9e-86;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGCTCTCCAGCAATCTCTCTGCATCTCCAGGGGAAAAGTCCACA 60  
DB 1 CAAATGTTCTCTCCAGCTCTCCAGCAATCTCTCTGCATCTCCAGGGGAAAAGTCCACA 60  
QY 61 ATGACTTGACAGGGGCGAGCTCAAGTAAATTAATCATGACAGGGTACAGCAGAGCCAGGA 120  
DB 61 ATGACTTGACAGGGGCGAGCTCAAGTAAATTAATCATGACAGGGTACAGCAGAGCCAGGA 120  
QY 121 TCCTCCCGCCAAACCCGTGATTTCTGCGACATCCACCTGGCTTCTGGAGTCCCTGCTCGC 180  
DB 121 TCCTCCCGCCAAACCCGTGATTTCTGCGACATCCACCTGGCTTCTGGAGTCCCTGCTCGC 180  
QY 181 TTCAGTGGCAGTGGGTCTGGAGACCTTCTCTCTCATCATCAGCAGAGTGGAGCTGAA 240  
DB 181 TTCAGTGGCAGTGGGTCTGGAGACCTTCTCTCTCATCATCAGCAGAGTGGAGCTGAA 240  
QY 241 GATGCTGCGCATTATTACTGCGCAGCAGTGGAGTAAACCCAGCCAGTTCGAGGGGG 300  
DB 241 GATGCTGCGCATTATTACTGCGCAGCAGTGGAGTAAACCCAGCCAGTTCGAGGGGG 300  
QY 301 ACCAGTCTGGAAATTAAGA 318  
DB 301 ACCAGTCTGGAAATTAAGA 318

RESULT 3  
US-08-149-099C-6  
Sequence 6, Application US/08149099C  
Patent No. 5736137

GENERAL INFORMATION:

APPLICANT: ANDERSON, Darrell R.  
APPLICANT: HANNA, Nabil  
APPLICANT: LEONARD, John E.  
APPLICANT: NEWMAN, Roland A.  
APPLICANT: REFF, Mitchell E.  
APPLICANT: RASTETTER, William H.  
TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND  
TITLE OF INVENTION: RADIOLABELLED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED  
TITLE OF INVENTION: DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P O Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/149,099C  
APPLICATION NUMBER: US/08/149,099C  
FILING DATE: 03-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/978,891  
FILING DATE: 12-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-014  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 384 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..384  
NAME/KEY: mat-peptide  
LOCATION: 67..384  
US-08-149-099C-6

Query Match 94.5%; Score 300.4; DB 1; Length 384;  
Best Local Similarity 96.5%; Pred. No. 9.7e-86;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGCTCCAGCAATCCTGTGATCTCCAGGGAAGGTCACA 60  
DB 67 CAAATGTTCTCTCCAGCTCCAGCAATCCTGTGATCTCCAGGGAAGGTCACA 126  
QY 61 ATGACTTGAGGAGCCAGCTCAAGTGTAAATTACATGCATGCTGACGAGCAAGCAGGA 120  
DB 127 ATGACTTGAGGAGCCAGCTCAAGTGTAAATTACATGCATGCTGACGAGCAAGCAGGA 186  
QY 121 TCCTCCGCCAAACCTGGATTTCTGCGACATCCAACTGGCTTCTGGAGTCCCTGCTGC 180  
DB 187 TCCTCCGCCAAACCTGGATTTCTGCGACATCCAACTGGCTTCTGGAGTCCCTGCTGC 246  
QY 181 TTCAGTGGCAGTGGGCTGGGACCTCTTACTCTCACAATCAGACAGAGTGGAGGCTGAA 240  
DB 247 TTCAGTGGCAGTGGGCTGGGACCTCTTACTCTCACAATCAGACAGAGTGGAGGCTGAA 306  
QY 241 GATGCTGCCACTTATTACTGCGCAGAGTGGAGTAGTAACCCACGCTTTCGAGGAGG 300  
DB 307 GATGCTGCCACTTATTACTGCGCAGAGTGGAGTAGTAACCCACGCTTTCGAGGAGG 366  
QY 301 ACCATGCTGGAATAAGA 318  
DB 367 ACCAAGCTGGAATCAAA 384

RESULT 4  
US-08-478-967A-6  
Sequence 6, Application US/08478967A  
Patent No. 5843439  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, Darrell R.  
APPLICANT: HANNA, Nabil  
APPLICANT: LEONARD, John E.  
APPLICANT: NEWMAN, Roland A.  
APPLICANT: REFF, Mitchell E.  
APPLICANT: RASTETTER, William H.

TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND  
TITLE OF INVENTION: RADIOLABELLED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED  
TITLE OF INVENTION: DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL  
LYMPHOMA  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22133-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,967A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
APPLICATION NUMBER: US 07/978,891  
FILING DATE: 12-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-014  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 384 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..384  
NAME/KEY: mat-peptide  
LOCATION: 67..384  
US-08-478-967A-6

Query Match 94.5%; Score 300.4; DB 2; Length 384;  
Best Local Similarity 96.5%; Pred. No. 9.7e-86;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGCTCCAGCAATCCTGTGATCTCCAGGGAAGGTCACA 60  
DB 67 CAAATGTTCTCTCCAGCTCCAGCAATCCTGTGATCTCCAGGGAAGGTCACA 126  
QY 61 ATGACTTGAGGAGCCAGCTCAAGTGTAAATTACATGCATGCTGACGAGCAAGCAGGA 120  
DB 127 ATGACTTGAGGAGCCAGCTCAAGTGTAAATTACATGCATGCTGACGAGCAAGCAGGA 186  
QY 121 TCCTCCGCCAAACCTGGATTTCTGCGACATCCAACTGGCTTCTGGAGTCCCTGCTGC 180  
DB 187 TCCTCCGCCAAACCTGGATTTCTGCGACATCCAACTGGCTTCTGGAGTCCCTGCTGC 246  
QY 181 TTCAGTGGCAGTGGGCTGGGACCTCTTACTCTCACAATCAGACAGAGTGGAGGCTGAA 240  
DB 247 TTCAGTGGCAGTGGGCTGGGACCTCTTACTCTCACAATCAGACAGAGTGGAGGCTGAA 306  
QY 241 GATGCTGCCACTTATTACTGCGCAGAGTGGAGTAGTAACCCACGCTTTCGAGGAGG 300  
DB 307 GATGCTGCCACTTATTACTGCGCAGAGTGGAGTAGTAACCCACGCTTTCGAGGAGG 366



Best Local Similarity 96.5%; Pred. No. 3.2e-85;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCCACTTCACCAATCTCTGTCATCTCCAGGGGAAAAGTCA 60  
Db 1045 CAAATGTTCTCTCCCACTTCACCAATCTCTGTCATCTCCAGGGGAAAAGTCA 1104  
QY 61 ATGACTGCAGGGCCAGCTCAAGTGTAAATTACATGCACGTGTACACAGAGCCAGGA 120  
Db 1105 ATGACTGCAGGGCCAGCTCAAGTGTAAATTACATGCACGTGTACACAGAGCCAGGA 1164  
QY 121 TCTCTCCCAAAACCTGGATTCTGCGACATCAACCTGGCTTGTGAGTCCCTGCTGC 180  
Db 1165 TCTCTCCCAAAACCTGGATTCTGCGACATCAACCTGGCTTGTGAGTCCCTGCTGC 1224  
QY 181 TTCAGTGCAGTGGGTGTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 240  
Db 1225 TTCAGTGCAGTGGGTGTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 1284  
QY 241 GATGCTGCCACTTATTACTGCCAGAGTGCAGTGTAACTAACCCACCTGGAGGGGG 300  
Db 1285 GATGCTGCCACTTATTACTGCCAGAGTGCAGTGTAACTAACCCACCTGGAGGGGG 1344  
QY 301 ACCATGCTGAAATPAGA 318  
Db 1345 ACCAAGCTGGAATCAAA 1362

RESULT 7  
US-08-478-967A-3  
Sequence 3, Application US/08478967A  
Patent No. 5843439

## GENERAL INFORMATION:

APPLICANT: ANDERSON, Darrell R.

APPLICANT: HANNA, Nabli

APPLICANT: LEONARD, John E.

APPLICANT: NEWMAN, Roland A.

APPLICANT: REFF, Mitchell E.

APPLICANT: RASTETTER, William H.

TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND

TITLE OF INVENTION: RADIOLABELLED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED

TITLE OF INVENTION: DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER &amp; MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478, 967A

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/149, 099

FILING DATE: 03-NOV-1993

APPLICATION NUMBER: US 07/978, 891

FILING DATE: 12-NOV-1992

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35, 030

REFERENCE/DOCKET NUMBER: 012712-014

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 3:

## SEQUENCE CHARACTERISTICS:

LENGTH: 9209 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHEITICAL: NO

ANTI-SENSE: NO

US-08-478-967A-3

Query Match

Best Local Similarity 94.5%; Score 300.4; DB 2; Length 9209;

Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCCACTTCACCAATCTCTGTCATCTCCAGGGGAAAAGTCA 60  
Db 1045 CAAATGTTCTCTCCCACTTCACCAATCTCTGTCATCTCCAGGGGAAAAGTCA 1104  
QY 61 ATGACTGCAGGGCCAGCTCAAGTGTAAATTACATGCACGTGTACACAGAGCCAGGA 120  
Db 1105 ATGACTGCAGGGCCAGCTCAAGTGTAAATTACATGCACGTGTACACAGAGCCAGGA 1164  
QY 121 TCTCTCCCAAAACCTGGATTCTGCGACATCAACCTGGCTTGTGAGTCCCTGCTGC 180  
Db 1165 TCTCTCCCAAAACCTGGATTCTGCGACATCAACCTGGCTTGTGAGTCCCTGCTGC 1224  
QY 181 TTCAGTGCAGTGGGTGTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 240  
Db 1225 TTCAGTGCAGTGGGTGTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGCTGAA 1284  
QY 241 GATGCTGCCACTTATTACTGCCAGAGTGCAGTGTAACTAACCCACCTGGAGGGGG 300  
Db 1285 GATGCTGCCACTTATTACTGCCAGAGTGCAGTGTAACTAACCCACCTGGAGGGGG 1344  
QY 301 ACCATGCTGAAATPAGA 318  
Db 1345 ACCAAGCTGGAATCAAA 1362

RESULT 8  
US-08-475-815B-3  
Sequence 3, Application US/08475815B  
Patent No. 639061

## GENERAL INFORMATION:

APPLICANT: ANDERSON, DARRELL R.

APPLICANT: HANNA, NABLI

APPLICANT: LEONARD, JOHN E.

APPLICANT: NEWMAN, ROLAND A.

APPLICANT: REFF, MITCHELL E.

APPLICANT: RASTETTER, WILLIAM H.

TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND

TITLE OF INVENTION: RADIOLABELLED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED

TITLE OF INVENTION: DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: PILLSBURY WINTHROP

STREET: 1100 New York Avenue, N.W., Ninth FL.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475, 815B

FILING DATE: 07-JUN-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/149, 099

```

1      FILING DATE: 03-NOV-1993
2      PRIOR APPLICATION DATA:
3      APPLICATION NUMBER: US 07/978,891
4      FILING DATE: 13-NOV-1992
5      ATTORNEY/AGENT INFORMATION:
6      NAME: Teskin, Robin L.
7      REGISTRATION NUMBER: 35,030
8      REFERENCE/DOCKET NUMBER: 23522-0157
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE: 202-861-3000
11     TELEFAX: 202-822-0944
12     INFORMATION FOR SEQ ID NO: 3:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 9209 base pairs
15     TYPE: nucleic acid
16     STRANDEDNESS: single
17     TOPOLOGY: linear
18     MOLECULE TYPE: DNA (genomic)
19     US-08-475-815B-3

```

Query Match	94.58;	Score 300.4;	DB 4;	Length 9209;
Best Local Similarity	96.58;	Pred. No. 3.2e-85;		
Matches 307; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0

QY	1	CAAATTTCTTCTCTCCAGATCCAGCAAAATCTGTGATCTTCAGAGGGGAAAGGTGACA	60
Db	1045	CAAAATTTCTCTCTCCAGATCCAGCAAAATCTGTGATCTTCAGAGGGGAAAGGTGACA	11045
QY	61	ATGACTTGCAGGGCCAGCTAACTGTAATTTACATGCACTGGTACGACAGCAAGCCAGAGA	120
Db	1105	ATGACTTGCAGGGCCAGCTAACTGTAATTTACATGCACTGGTACGACAGCAAGCCAGAGA	11645
QY	121	TCTCTCCCCCAAAACCCTGGATTTCTGGCACATCCAAACCTGGCTTCGAGATCCCTCTCGC	180
Db	1165	TCTCTCCCCCAAAACCCTGGATTTCTGGCACATCCAAACCTGGCTTCGAGATCCCTCTCGC	12245
QY	181	TTTCAGTGGCAGATGGGTGTGGGACCTCTTACTCTCTTACAAATCAGCAGATGGAGGCTGAA	240
Db	1225	TTTCAGTGGCAGATGGGTGTGGGACCTCTTACTCTCTTACAAATCAGCAGATGGAGGCTGAA	12845
QY	241	GATGCTGCGCACTTATTTACTCCAGCAGATGAGTAGTAACCCACCCTTTCGAGAGGGGG	300
Db	1285	GATGCTGCGCACTTATTTACTCCAGCAGATGAGTAGTAACCCACCCTTTCGAGAGGGGG	13445
QY	301	ACCATGCTGGAAATTAAGA 318	
Db	1345	ACCATGCTGGAAATTAAGA 1362	

RESULT 9  
 US-08-819-866-2  
 Sequence 2, Application US/08819866  
 Patent No. 5830698  
 GENERAL INFORMATION:  
 APPLICANT: REFF, Mitchell E.  
 APPLICANT: BARNETT, Richard Spence  
 APPLICANT: MCLACHLAN, Karen Retta  
 TITLE OF INVENTION: NOVEL METHOD FOR INTEGRATING GENES AT  
 TITLE OF INVENTION: SPECIFIC SITES IN MAMMALIAN CELLS VIA HOMOLOGOUS  
 RECOMBINATION AND VECTORS FOR ACCOMPLISHING THE SAME  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/819,866
3 FILING DATE: 14-MAR-1997
4 CLASSIFICATION: 435
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Teskin, Robin L.
7 REGISTRATION NUMBER: 35,030
8 REFERENCE/DOCKET NUMBER: 012712-352
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: (703) 836-6620
11 TELEFAX: (703) 836-2021
12 INFORMATION FOR SEQ ID NO: 2:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 18986 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: DNA (genomic)
19 US-08-819-866-2

```

Query Match	94.5%;	Score 300.4;	DB 2;	length 18986;
Best Local Similarity	96.5%;	Pred. No. 4.2e-85;		
Matches 307; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0

QY	1	CAAATGCTTCTCTCCAGAGTCTCCAGCAATCTGTGTGCAATTCACAGGGGAAAAGGTACCA	60
Db	8077	CAAAATGCTCTCTCCAGTCTCCAGCAATCTGTGTGCAATTCACAGGGGAAAAGGTACCA	8136
QY	61	ATGCACTTGCAAGGCGCAGCCCAAGTGTAAATTCATGCAATCTGGTACCACAGAAAGCCACGA	120
Db	8137	ATGCACTTGCAAGGCGCAGCTCAAGTGTAAATTCATGCAATCTGGTTCACAGCAAGCCACGA	8196
QY	121	TCTCTCCCCCAACCCCTGGATTTCTGTGCACATCCACAACTGGCTTGTGAGTCCCTGCTGC	180
Db	8197	TCTCTCCCCCAACCCCTGGATTTATGCCACATCCAACTGGCTTGTGAGTCCCTGCTGC	8256
QY	181	TTTCAGTGGCAGTGGGCTGTGGGACCTCTTACTCTCTACAAATGACAGAGTGGAGGCTAA	240
Db	8257	TTTCAGTGGCAGTGGGCTGTGGGACCTTCTTACTCTCTACATAGCAGAGTGGAGGCTAA	8316
QY	241	GATGCTGTCCACTATTAATCTCCAGCAGAGTGGAGTAGTAACCAACCACCTTGTGGAGGGGG	300
Db	8317	GATGCTGTCCACTATTAATCTCCAGCAGAGTGGAGTAGTAACCAACCACCTTGTGGAGGGGG	8376
QY	301	ACCATGCTGGAAATTAAGA	318
Db	8377	ACCAAGCTGGAAATTAAGA	8394

RESULT 10  
US-09-023-715-2  
Sequence 2, Application US/09023715  
Patent No. 5998144  
GENERAL INFORMATION:  
APPLICANT: REEF, Mitchell E.  
APPLICANT: BARNETT, Richard Spence  
APPLICANT: MCLACHLAN, Karen Reta  
TITLE OF INVENTION: NOVEL METHOD FOR INTEGRATING GENES AT  
TITLE OF INVENTION: SPECIFIC SITES IN MAMMALIAN CELLS VIA HOMOLOGOUS  
TITLE OF INVENTION: RECOMBINATION AND VECTORS FOR ACCOMPLISHING THE SAME  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
STREET: P.O. Box 1404  
City: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,715  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/819,866  
FILING DATE: 14-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-352  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18986 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-023-715-2

Query Match 94.5%; Score 300.4; DB 2; Length 18986;  
Best Local Similarity 96.5%; Pred. No. 4,2e-85;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGCTCCAGCAATCTGTCTGCATCTCCAGGGAAGGTCACA 60  
DB 8077 CAAATGTTCTCTCCAGCTCCAGCAATCTGTCTGCATCTCCAGGGAAGGTCACA 8136  
QY 61 ATGACTTGACAGGCGCAGCTCAAGTAAATACATGACCTGTACACGACAGCAAGCAGGA 120  
DB 8137 ATGACTTGACAGGCGCAGCTCAAGTAAATACATGACCTGTACACGACAGCAAGCAGGA 8196  
QY 121 TCCTCCCCCAAAACCCTGGATTTCGCCACATCAACCTGGCTTGGAGTCCCTGCTGC 180  
DB 8197 TCCTCCCCCAAAACCCTGGATTTCGCCACATCAACCTGGCTTGGAGTCCCTGCTGC 8256  
QY 181 TTCAGTGGACAGTGGGTCTGGAGCTCTTACTCTCTCACAATCAGACAGAGTGGAGCTGAA 240  
DB 8257 TTCAGTGGACAGTGGGTCTGGAGCTCTTACTCTCTCACAATCAGACAGAGTGGAGCTGAA 8316  
QY 241 GATCTGCGCACTTATTAATGCGCAGAGTGAAGTAAACCCACGCTTGGAGGGGG 300  
DB 8317 GATCTGCGCACTTATTAATGCGCAGAGTGAAGTAAACCCACGCTTGGAGGGGG 8376  
QY 301 ACCATGCTGGAATAGA 318  
DB 8377 ACCAAGCTGGAAATCAAA 8394

RESULT 11  
US-09-343-485A-2  
Sequence 2, Application US/09343485A  
Patent No. 6413777  
GENERAL INFORMATION:  
APPLICANT: REEF, MITCHELL R.  
APPLICANT: BARRETT, RICHARD S.  
APPLICANT: MCILACHLAN, KAREN R.  
TITLE OF INVENTION: NOVEL METHOD FOR INTEGRATING GENES AT SPECIFIC SITES IN  
TITLE OF INVENTION: MAMMALIAN CELLS VIA HOMOLOGOUS RECOMBINATION AND  
TITLE OF INVENTION: VECTORS FOR ACCOMPLISHING THE SAME  
FILE REFERENCE: 037003-0275807  
CURRENT APPLICATION NUMBER: US/09/343,485A  
CURRENT FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: 09/023,715  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: 08/819,866  
PRIOR FILING DATE: 1997-03-14  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2

LENGTH: 18986  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
OTHER INFORMATION: referred to as "Molly"  
US-09-343-485A-2

Query Match 94.5%; Score 300.4; DB 4; Length 18986;  
Best Local Similarity 96.5%; Pred. No. 4,2e-85;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGCTCCAGCAATCTGTCTGCATCTCCAGGGAAGGTCACA 60  
DB 8077 CAAATGTTCTCTCCAGCTCCAGCAATCTGTCTGCATCTCCAGGGAAGGTCACA 8136  
QY 61 ATGACTTGACAGGCGCAGCTCAAGTAAATACATGACCTGTACACGACAGCAAGCAGGA 120  
DB 8137 ATGACTTGACAGGCGCAGCTCAAGTAAATACATGACCTGTACACGACAGCAAGCAGGA 8196  
QY 121 TCCTCCCCCAAAACCCTGGATTTCGCCACATCAACCTGGCTTGGAGTCCCTGCTGC 180  
DB 8197 TCCTCCCCCAAAACCCTGGATTTCGCCACATCAACCTGGCTTGGAGTCCCTGCTGC 8256  
QY 181 TTCAGTGGACAGTGGGTCTGGAGCTCTTACTCTCTCACAATCAGACAGTGGAGCTGAA 240  
DB 8257 TTCAGTGGACAGTGGGTCTGGAGCTCTTACTCTCTCACAATCAGACAGTGGAGCTGAA 8316  
QY 241 GATCTGCGCACTTATTAATGCGCAGAGTGAAGTAAACCCACGCTTGGAGGGGG 300  
DB 8317 GATCTGCGCACTTATTAATGCGCAGAGTGAAGTAAACCCACGCTTGGAGGGGG 8376  
QY 301 ACCATGCTGGAATAGA 318  
DB 8377 ACCAAGCTGGAAATCAAA 8394

RESULT 12  
US-08-476-275-3  
Sequence 3, Application US/08476275  
Patent No. 5776456  
GENERAL INFORMATION:  
APPLICANT: Anderson, Darrell R.  
APPLICANT: Hanna, Nabil  
APPLICANT: Leonard, John E.  
APPLICANT: Newman, Roland A.  
APPLICANT: Reiff, Mitchell E.  
APPLICANT: Rastetter, William H.  
TITLE OF INVENTION: Therapeutic Application of Chimeric and  
TITLE OF INVENTION: Radiolabeled Antibodies to Human B Lymphocyte Restricted  
TITLE OF INVENTION: Differentiation Antigen for the Treatment of B-cell  
Lymphoma  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince St.  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,275  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/978,891  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-155  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 384 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: murine variable region light chain  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..384  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 67..384  
US-08-476-275-3

Query Match 94.0%; Score 298.8; DB 1; Length 384;  
Best Local Similarity 96.2%; Pred. No. 3.1e-85;  
Matches 306; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGTCCTCCAGCAATCTCTGATCTTCAGGGGAAAGTCCACA 60  
DB 67 CAAATGTTCTCTCCAGTCCTCCAGCAATCTCTGATCTTCAGGGGAAAGTCCACA 126  
QY 61 ATGACTTGACGGGCGGCTCAAGTGAATTAATATGACACTGGTACAGAGAGCCAGCA 120  
DB 127 ATGACTTGACGGGCGGCTCAAGTGAATTAATATGACACTGGTACAGAGAGCCAGCA 186  
QY 121 TCCTCCCGCCAAACCTGATTTCTGCCACATCCAACTGGCTTCTGAGTCCCTGCTCG 180  
DB 187 TCCTCCCGCCAAACCTGATTTCTGCCACATCCAACTGGCTTCTGAGTCCCTGCTCG 246  
QY 181 TTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240  
DB 247 TTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 306  
QY 241 GATGCTGCCACTTATTAAGTCCAGCAGTGGAGTAAACCCAGCTTCGGAGGGGG 300  
DB 307 GATGCTGCCACTTATTAAGTCCAGCAGTGGAGTAAACCCAGCTTCGGAGGGGG 366  
QY 301 ACCAGCTGGAAATAGA 318  
DB 367 ACCAGCTGGAAATAGA 384

RESULT 13  
US-08-476-275-2

Sequence 2, Application US/08476275  
Patent No. 5776456  
GENERAL INFORMATION:  
APPLICANT: Anderson, Darrell R.  
APPLICANT: Hanna, Nabil  
APPLICANT: Leonard, John E.  
APPLICANT: Newman, Roland A.  
APPLICANT: Reff, Mitchell E.  
APPLICANT: Rastetter, William H.  
TITLE OF INVENTION: Therapeutic Application of Chimeric and  
TITLE OF INVENTION: Radiolabeled Antibodies to Human B Lymphocyte Restricted  
TITLE OF INVENTION: Differentiation Antigen for the Treatment of B-Cell  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: 699 Prince St.  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,275  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/978,891  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-155  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9209 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: anti-CD20 in TCAE 8  
US-08-476-275-2

Query Match 93.5%; Score 297.2; DB 1; Length 9209;  
Best Local Similarity 95.9%; Pred. No. 3.3e-84;  
Matches 305; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGTCCTCCAGCAATCTCTGATCTTCAGGGGAAAGTCCACA 60  
DB 1045 CAAATGTTCTCTCCAGTCCTCCAGCAATCTCTGATCTTCAGGGGAAAGTCCACA 1104  
QY 61 ATGACTTGACGGGCGGCTCAAGTGAATTAATATGACACTGGTACAGAGAGCCAGCA 120  
DB 1105 ATGACTTGACGGGCGGCTCAAGTGAATTAATATGACACTGGTACAGAGAGCCAGCA 1164  
QY 121 TCCTCCCGCCAAACCTGATTTCTGCCACATCCAACTGGCTTCTGAGTCCCTGCTCG 180  
DB 1165 TCCTCCCGCCAAACCTGATTTCTGCCACATCCAACTGGCTTCTGAGTCCCTGCTCG 1224  
QY 181 TTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240  
DB 1225 TTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 1284  
QY 241 GATGCTGCCACTTATTAAGTCCAGCAGTGGAGTAAACCCAGCTTCGGAGGGGG 300  
DB 1285 GATGCTGCCACTTATTAAGTCCAGCAGTGGAGTAAACCCAGCTTCGGAGGGGG 1344  
QY 301 ACCAGCTGGAAATAGA 318  
DB 1345 ACCAGCTGGAAATAGA 1362

RESULT 14

US-08-783-853A-104  
Sequence 104, Application US/08783853A  
Patent No. 6005091



GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Giora  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
TITLE OF INVENTION: OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/783,853A  
FILING DATE: 16-JAN-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/029,119  
FILING DATE: 24-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1..318  
OTHER INFORMATION:  
US-08-783-853A-104

Query Match 93.0%; Score 295.6; DB 3; Length 318;  
Best Local Similarity 95.6%; Pred.No.2.9e-84;  
Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CAAATGTTCTCTCCAGTCTCCAGCAATCGTGTGATCTCCAGGAGAAAGTGCACA 60  
DB 1 CAGATAGTACTCTCCAGTCTCCAGCAATCGTGTGATCTCCAGGAGAAAGTGCACA 60  
QY 61 ATGACTTGCAGGCGCAGCTCAAGTGTAAATACATGCACTGGTACCAGAGAAGCCAGGA 120  
DB 61 ATGACTTGCAGGCGCAGCTCAAGTGTAAATACATGCACTGGTACCAGAGAAGCCAGGA 120  
QY 121 TCCTCCGCCAAACCCGATGATTTCTGCGACATCCAACTGGCTTTGAGATCCCTGCTCGC 180  
DB 121 TCCTCCGCCAAACCCGATGATTTCTGCGACATCCAACTGGCTTTGAGATCCCTGCTCGC 180  
QY 181 TTCAGTGCAGTGGTCTGTGGACCTCTTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240

|||||  
DB 181 TTCAGTGCAGTGGTCTGTGGACCTCTTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240  
QY 241 GATGCTGCCACTTATTACTGCCAGAGTGGAGTAGTAAACCCACCCAGTTCCGAGGGGG 300  
DB 241 GATGCTGCCACTTATTACTGCCAGAGTGGAGTAGTAAACCCACCCAGTTCCGAGGGGG 300  
QY 301 ACCATGCTGGAATAAGA 318  
DB 301 ACCATGCTGGAATCAAA 318

RESULT 15  
US-09-344-050-104  
Sequence 104 Application US/09344050  
Patent No. 6391289

GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Giora  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
TITLE OF INVENTION: OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,050  
FILING DATE: 24-JUN-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1..318  
OTHER INFORMATION:  
US-09-344-050-104

Query Match 93.0%; Score 295.6; DB 4; Length 318;



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OW protein - protein search, using sw model

Run on: November 27, 2002, 05:38:32 : Search time 58.5251 Seconds

(without alignments)  
241.342 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 556

Sequence: 1 QIVLSQSPAILASPGKVT.....CQKMSNPFGGTMLEIR 106

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_101002: \*  
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	526	94.6	106	AAW24532	Anti-Factor IX Mab
2	526	94.6	106	AAU81002	Mouse-human light
3	526	94.6	107	AAW24520	Mouse anti-human F
4	526	94.6	107	AAU80976	Murine BC2 light c
5	526	94.6	112	AAW24531	Anti-Factor IX Mab
6	526	94.6	112	AAU81001	Murine BC2 light c
7	521	93.7	128	AAW5214	Murine variable re
8	517	93.0	412	AAW30694	A fusion of anti-C
9	517	93.0	423	AAW30695	A fusion of anti-C
10	516	92.8	128	AAW70628	Sequence encoded b

11	516	92.8	128	10	AAW94781
12	516	92.8	128	18	AAW10589
13	516	92.8	128	18	AAW16344
14	516	92.8	128	18	AAW10243
15	516	92.8	128	18	AAW47521
16	516	92.8	128	19	AAW47514
17	516	92.8	128	19	AAW47514
18	516	92.8	128	20	AAW89541
19	516	92.8	128	22	AAW89093
20	514	92.4	633	21	AAW84965
21	507	91.2	129	13	AAW20789
22	507	91.2	129	13	AAW30430
23	507	91.2	235	17	AAW06178
24	507	91.2	235	20	AAW82746
25	507	91.2	235	20	AAW82746
26	507	91.2	235	21	AAW80625
27	504	90.6	145	12	AAW15323
28	504	90.6	145	14	AAW32127
29	503	90.5	106	23	AAW72834
30	501	90.1	105	18	AAW19914
31	491	88.3	108	16	AAW79883
32	489	87.9	236	15	AAW45442
33	489	87.9	237	17	AAW45442
34	489	87.9	237	18	AAW15185
35	486	87.4	107	14	AAW34263
36	486	87.4	246	15	AAW60522
37	484	87.1	252	21	AAW12562
38	484	87.1	252	22	AAW36825
39	482	86.7	107	13	AAW79059
40	478	86.0	107	13	AAW24723
41	475	85.4	97	11	AAW07317
42	475	85.4	97	17	AAW14485
43	475	85.4	97	17	AAW98874
44	475	85.4	97	20	AAW34426
45	475	85.4	97	21	AAW51369

## ALIGNMENTS

RESULT 1  
AAW24532 standard; Protein: 106 AA.  
AAW24532:  
27-DEC-1997 (first entry)  
Anti-Factor IX Mab chimeric light chain.  
Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;  
Chimeric antibody; antibody engineering; light chain.  
Chimeric Mus musculus.  
Chimeric Homo sapiens.  
Chimeric synthetic.  
WO9726010-A1.  
24-JUL-1997.  
17-JAN-1997; 97WO-US00759.  
24-OCT-1996; 96US-0029119.  
17-JAN-1996; 96US-0010108.  
(SMK) SMITHKLINE BEECHAM CORP.  
(UYVE) UNIV VERMONT & STATE AGRIC COLLEGE.  
Blackburn MN, Church WR, Feuerstein GZ, Gross MS;  
Nichols AJ, Padlan EA, Patel AH, Sylvester CR;  
WPI, 1997-385117/35.

2 H7 VL gene. Syn  
2H7 antibody light  
2H7 light chain va  
light chain variab  
Mouse 2H7 antibody  
Mouse 2H7 antibody  
Mouse 2H7 antibody  
Mouse antibody 2H7  
2H7 light chain va  
Amino acid sequenc  
Unprocessed variab  
ASB7 antibody light  
Murine ASB7 light  
Plasmid pNC3/ASB7  
A dimeric anti-CD2  
Anti-IL2R-alpha an  
Anti-NG2D hybrido  
Kappa light chain  
Anti-EGFR antibody  
Sequence of the si  
Anti-erbB2 scfv.  
Single-chain anti-  
Mab32 V-gene heavy  
Anti-TNF antibody D  
Chimeric L6 anti-t  
L6 scv protein. U  
Amino acid sequenc  
Sequence encoded b  
VK domain of antib  
Monoclonal antibod  
Monoclonal antibod  
Mouse monoclonal a

DR N-PSDB; AAT79900.  
 XX Inhibiting thrombosis with self-limiting antibody to coagulation  
 PT factor - avoids uncontrolled bleeding by providing only partial  
 PT inhibition  
 XX Example 7; Page 128; 150pp; English.  
 PS  
 CC This polypeptide comprises a mouse-human chimeric antibody  
 CC light chain in which the variable region is derived from mouse  
 CC anti-human factor IX monoclonal antibody BC2 (see AAW24531) and  
 CC human sequences from the immunoglobulin RF-TS3/CL framework.  
 CC It can be expressed in transfected mammalian cells utilizing a  
 CC cDNA construct (see AAT79900) obtained by PCR amplification (see  
 CC AAT79897-98) of BC2 cDNA and insertion of the PCR product into  
 CC F9H2HC 1-3 cDNA (see AAT77374). Claimed anti-Factor IX chimeric  
 CC antibodies are useful in the treatment of thrombosis.  
 XX  
 SO Sequence 106 AA:  
 Query Match 94.6%; Score 526; DB 18; Length 106;  
 Best Local Similarity 95.3%; Pred. No. 1.8e-30;  
 Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 QIVLSQSPALLSAPGKRYMTTCRASSSVNMYHQKPPSSPKPMISATSNLASSVPAR 60  
 DB 1 QIVLSQSPALLSAPGKRYMTTCRASSSVNMYHQKPPSSPKPMISATSNLASSVPAR 60  
 QY 61 FSGSGSGTSTLTSTRVEAEDAATYYCQOWSSNPPTFGGTMLEIR 106  
 DB 61 FSGSGSGTSTLTSTRVEAEDAATYYCQOWSSNPPTFGGTMLEIR 106  
 Db 61 FSGSGSGTSTLTSTRVEAEDAATYYCQOWSSNPPTFGGTMLEIR 106  
 RESULT 2  
 ID AAW81002 standard; Protein; 106 AA.  
 XX  
 AC AAW81002:  
 DT 09-APR-2002 (first entry)  
 XX  
 DE Mouse-human light chain polypeptide.  
 XX  
 KW Human: mouse; BC2; animal post-thromboembolic induced ischaemia;  
 KW thrombolytic agent; anti-Factor IX antibody; plasminogen activator;  
 KW thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;  
 KW vasotrophic; cardiant; anti-respiratory syncytial virus;  
 KW heavy chain variable region; light chain variable region.  
 XX  
 OS Chimeric - Mus sp.  
 OS Chimeric - Homo sapiens.  
 XX  
 PN WO200187339-A1.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 05-OCT-2000; 2000WO-US27438.  
 XX  
 PR 15-MAY-2000; 2000US-0571434.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR,  
 DR WPI; 2002-082944/11.  
 DR N-PSDB; ABR24005.  
 XX  
 XX Treating post-thromboembolic induced ischaemia in an animal by  
 PT administering anti-factor IX antibody in combination with a plasminogen  
 PT activator  
 XX  
 PS Example 7; Page 155-156; 163pp; English.  
 XX

CC The invention relates to a method for treating an animal  
 CC post-thromboembolic induced ischaemia or reducing a required dose of a  
 CC thrombolytic agent in treatment of an animal post-thromboembolic induced  
 CC ischaemia, comprising administering an anti-factor IX antibody or its  
 CC fragment, optionally in combination with a plasminogen activator or  
 CC thrombolytic agent. The method is useful for treating thromboembolic  
 CC post-thromboembolic-induced ischaemia, for preventing thromboembolic  
 CC stroke in an animal, and for reducing a required dose of a thrombolytic  
 CC agent. Sequences AAW80972-AAW81004 represent antibodies and vector  
 CC polypeptides used in the method of the invention.  
 XX  
 SO Sequence 106 AA:  
 Query Match 94.6%; Score 526; DB 23; Length 106;  
 Best Local Similarity 95.3%; Pred. No. 1.8e-30;  
 Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 QIVLSQSPALLSAPGKRYMTTCRASSSVNMYHQKPPSSPKPMISATSNLASSVPAR 60  
 DB 1 QIVLSQSPALLSAPGKRYMTTCRASSSVNMYHQKPPSSPKPMISATSNLASSVPAR 60  
 QY 61 FSGSGSGTSTLTSTRVEAEDAATYYCQOWSSNPPTFGGTMLEIR 106  
 DB 61 FSGSGSGTSTLTSTRVEAEDAATYYCQOWSSNPPTFGGTMLEIR 106  
 Db 61 FSGSGSGTSTLTSTRVEAEDAATYYCQOWSSNPPTFGGTMLEIR 106  
 RESULT 3  
 ID AAW24520 standard; Protein; 107 AA.  
 XX  
 AC AAW24520:  
 DT 26-DEC-1997 (first entry)  
 XX  
 DE Mouse anti-human Factor IX antibody BC2 light chain variable region.  
 XX  
 KW Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;  
 KW humanised antibody; antibody engineering; light chain; CDR;  
 KW complementarity determining region; myocardial infarction;  
 KW angina; atrial fibrillation; stroke; kidney damage;  
 KW pulmonary embolism; deep vein thrombosis; coronary angioplasty;  
 KW disseminated intravascular coagulation; artificial organ; sepsis;  
 KW shunt; prosthesis.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 24..33  
 FT /label= CDR1  
 FT /note= "(Claim 23)"  
 FT Region 49..55  
 FT /label= CDR2  
 FT /note= "(Claim 23)"  
 FT Region 88..96  
 FT /label= CDR3  
 FT /note= "(Claim 23)"  
 WO9726010-A1.  
 PD 24-JUL-1997.  
 XX  
 PF 17-JAN-1997; 97WO-US00759.  
 XX  
 PR 24-OCT-1996; 96US-0029119.  
 XX  
 PR 17-JAN-1996; 96US-0010108.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.  
 PI Blackburn MN, Church WR, Feuerstein GZ, Gross MS;  
 PI Nichols AJ, Padlan EA, Patel AH, Sylvester DR;  
 DR WPI; 1997-385117/35.  
 XX

DR N-PSDB; AAT77377.  
 XX Inhibiting thrombosis with self-limiting antibody to coagulation  
 PT factor - avoids uncontrolled bleeding by providing only partial  
 PT inhibition  
 XX  
 PS Example 5; Page 66-67; 150pp; English.  
 XX  
 CC This sequence comprises the light chain variable region of mouse  
 CC anti-human factor IX monoclonal antibody BC2. Claimed humanised  
 CC antibodies (see AAM24510-18) contain CDRs (see AAM24504-09) of BC2  
 CC heavy and light chain variable regions inserted into framework  
 CC regions of selected human antibody sequences. They have self-  
 CC limiting neutralising activity, and are useful as anticoagulant  
 CC agents in treatment of thrombosis associated with myocardial  
 CC infarction, unstable angina, atrial fibrillation, stroke, renal  
 CC damage, pulmonary embolism, deep vein thrombosis, percutaneous  
 CC transluminal coronary angioplasty, disseminated intravascular  
 CC coagulation, sepsis, or artificial organs, shunts or prostheses  
 CC (claimed). Also claimed are chimeric antibodies (see AAM24532), and  
 CC Fab and F(ab')<sub>2</sub> fragments. The claimed antibodies do not cause  
 CC uncontrolled bleeding (contrast heparin and warfarin) since they  
 CC provide only partial inhibition of coagulation.  
 XX  
 SQ Sequence 107 AA:  
 Query Match 94.6%; Score 526; DB 18; Length 107;  
 Best Local Similarity 95.3%; Pred. No. 1.8e-30;  
 Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 OY 1 QIVLSQSPAILLSAPGKVTMTCRASSSVNMYHWYQOKPGSSPKPIVATSNLASGVPAR 60  
 DB 1 QIVLSQSPAILLSAPGKVTMTCRASSSVNMYHWYQOKPGSSPKPIVATSNLASGVPAR 60  
 OY 61 FSGSGSGTSYSLTISRVEAEDATYCCQWSSNPPTFGGTMLEIR 106  
 DB 61 FSGSGSGTSYSLTISRVEAEDATYCCQWSSNPPTFGGTMLEIR 106  
 RESULT 4  
 AAU80976  
 ID AAU80976 standard; Protein: 107 AA.  
 XX  
 AC AAU80976;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE Murine BC2 light chain variable region.  
 XX  
 KM Human; mouse; BC2; animal post-thromboembolic induced ischaemia;  
 KM thrombolytic agent; anti-factor IX antibody; plasminogen activator;  
 KM thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;  
 KM vasotropic; cardiant; anti-respiratory syncytial virus;  
 KM heavy chain variable region; light chain variable region.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200187339-A1.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 05-OCT-2000; 2000MO-US27438.  
 XX  
 PR 15-MAY-2000; 2000US-0571434.  
 XX  
 PA (SMK ) SMITHKLINE BEECHAM CORP.  
 XX  
 PI Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;  
 XX  
 DR WPI; 2002-082944/11.  
 XX  
 DR N-PSDB; ABR23937.  
 XX  
 PT Treating post-thromboembolic induced ischaemia in an animal by

PT administering anti-factor IX antibody in combination with a plasminogen  
 PT activator  
 XX  
 XX Example 5; Page 97; 163pp; English.  
 PS  
 XX  
 CC The invention relates to a method for treating an animal  
 CC post-thromboembolic induced ischaemia or reducing a required dose of a  
 CC thrombolytic agent in treatment of an animal post-thromboembolic induced  
 CC ischaemia, comprising administering an anti-factor IX antibody or its  
 CC fragment, optionally in combination with a plasminogen activator or  
 CC thrombolytic agent. The method is useful for treating  
 CC post-thromboembolic-induced ischaemia, for preventing thromboembolic  
 CC stroke in an animal, and for reducing a required dose of a thrombolytic  
 CC agent. Sequences AAU80972-AAU81004 represent antibodies and vector  
 CC polypeptides used in the method of the invention.  
 XX  
 SQ Sequence 107 AA:  
 Query Match 94.6%; Score 526; DB 23; Length 107;  
 Best Local Similarity 95.3%; Pred. No. 1.8e-30;  
 Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 OY 1 QIVLSQSPAILLSAPGKVTMTCRASSSVNMYHWYQOKPGSSPKPIVATSNLASGVPAR 60  
 DB 1 QIVLSQSPAILLSAPGKVTMTCRASSSVNMYHWYQOKPGSSPKPIVATSNLASGVPAR 60  
 OY 61 FSGSGSGTSYSLTISRVEAEDATYCCQWSSNPPTFGGTMLEIR 106  
 DB 61 FSGSGSGTSYSLTISRVEAEDATYCCQWSSNPPTFGGTMLEIR 106  
 RESULT 5  
 AAM24531  
 ID AAM24531 standard; Protein: 112 AA.  
 XX  
 AC AAM24531;  
 XX  
 DT 27-DEC-1997 (first entry)  
 XX  
 DE Anti-Factor IX Mab BC2 modified light chain.  
 XX  
 KM Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;  
 KM chimeric antibody; antibody engineering; light chain.  
 XX  
 OS Chimeric Mus musculus.  
 XX  
 OS Chimeric synthetic.  
 XX  
 PN WO9726010-A1.  
 XX  
 PD 24-JUL-1997.  
 XX  
 PF 17-JAN-1997; 97WO-US00759.  
 XX  
 PR 24-OCT-1996; 96US-0029119.  
 XX  
 PR 17-JAN-1996; 96US-0010108.  
 XX  
 PA (SMK ) SMITHKLINE BEECHAM CORP.  
 XX  
 PI (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.  
 XX  
 PI Blackburn MN, Church WR, Feuerstein GZ, Gross MS;  
 XX  
 PI Nichols AJ, Padlan EA, Patel AH, Sylvester DR;  
 XX  
 DR WPI; 1997-385117/35.  
 XX  
 DR N-PSDB; AAT79899.  
 XX  
 PT Inhibiting thrombosis with self-limiting antibody to coagulation  
 PT factor - avoids uncontrolled bleeding by providing only partial  
 PT inhibition  
 XX  
 PS Example 7; Page 126; 150pp; English.  
 XX  
 CC This polypeptide sequence comprises a modified light chain variable  
 CC region (see also AAM24520) of mouse anti-human factor IX monoclonal

CC antibody BC2. It is encoded by a DNA fragment (see AAT79899)  
CC produced by PCR amplification of BC2 cDNA. The modified BC2 VL  
CC region is incorporated into novel mouse-human chimeric light chain  
CC FgChIC (see AAM24532). Claimed anti-Factor IX chimeric antibodies  
CC are useful in the treatment of thrombosis.

XX Sequence 112 AA:

Query Match 94.6%; Score 526; DB 18; Length 112;  
Best Local Similarity 95.3%; Pred. No. 1.9e-30;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 QIVLSQSPAILASPGKVTMTCRASSSVNYMHYQOKPGSSPKPWISATSNLASGVPAR 60

DB 1 QIVLSQSPAILASPGKVTMTCRASSSVNYMHYQOKPGSSPKPWIVATSNLASGVPAR 60

OY 61 FSGSGSGTSTLTISRVEADATYCCQWSSNPPTFGGTMLEIR 106  
|||||  
DB 61 FSGSGSGTSTLTISRVEADATYCCQWSSNPPTFGGTMLEIR 106

#### RESULT 6

AAU81001 standard; Protein: 112 AA.

AC AAU81001;

DT 09-APR-2002 (first entry)

DE Murine BC2 light chain modified variable region.

Human; mouse; BC2; animal post-thromboembolic induced ischaemia;  
thrombolytic agent; anti-factor IX antibody; plasminogen activator;  
thromboembolic stroke; cerebroprotective; anticoagulant; thrombolytic;  
vasotropic; cardiac; anti-respiratory syncytial virus;  
heavy chain variable region; light chain variable region.

OS Mus sp.  
XX Synthetic.

PN WO200187339-A1.

PD 22-NOV-2001.

PF 05-OCT-2000; 2000WO-US27438.

PR 15-MAY-2000; 2000US-0571434.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;

DR WPI: 2002-082944/11.

DR N-PSDB; ABK24004.

PT Treating post-thromboembolic induced ischaemia in an animal by  
administering anti-factor IX antibody in combination with a plasminogen  
activator

XX Example 7; Page 153-154; 163pp; English.

The invention relates to a method for treating an animal  
post-thromboembolic induced ischaemia or reducing a required dose of a  
thrombolytic agent in treatment of an animal post-thromboembolic induced  
ischaemia, comprising administering an anti-factor IX antibody or its  
fragment, optionally in combination with a plasminogen activator or  
thrombolytic agent. The method is useful for treating thromboembolic  
post-thromboembolic-induced ischaemia, for preventing thromboembolic  
stroke in an animal, and for reducing a required dose of a thrombolytic  
agent. Sequences AAU80972-AAU81004 represent antibodies and vector  
polypeptides used in the method of the invention.

Sequence 112 AA;

Query Match 94.6%; Score 526; DB 23; Length 112;  
Best Local Similarity 95.3%; Pred. No. 1.9e-30;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 QIVLSQSPAILASPGKVTMTCRASSSVNYMHYQOKPGSSPKPWISATSNLASGVPAR 60

DB 1 QIVLSQSPAILASPGKVTMTCRASSSVNYMHYQOKPGSSPKPWIVATSNLASGVPAR 60

OY 61 FSGSGSGTSTLTISRVEADATYCCQWSSNPPTFGGTMLEIR 106  
|||||  
DB 61 FSGSGSGTSTLTISRVEADATYCCQWSSNPPTFGGTMLEIR 106

#### RESULT 7

AAU55214 standard; Protein: 128 AA.

AC AAU55214;

DT 01-FEB-1995 (first entry)

DE Murine variable region light chain from 2B5.

B cell lymphoma chimeric antibody; CD20; peripheral blood cells;  
cell lysis.

OS Mus musculus.

PN WO9411026-A.

PD 26-MAY-1994.

PF 12-NOV-1993; 93WO-US10953.

PR 13-NOV-1992; 92US-0978891.

PR 03-NOV-1993; 93US-0149099.

PA (IDEC-) IDEC PHARM CORP.

PI Anderson DR, Hanna N, Leonard JE, Newman RA, Rusterliet WH;

PI RefTime;

DR N-PSDB; Q65630.

DR WPI: 1994-183162/22.

PT Treating B cell lymphoma with chimeric antibody - against CD20,  
causing rapid depletion of peripheral B cells, also new  
antibodies and hybridomas

PS Disclosure; Fig 4; 101pp; English.

CC The sequence is the murine variable region light chain derived from

CC murine anti-CD20 monoclonal antibody 2B5.

CC See also AAQ65629-35.

XX Sequence 128 AA;

Query Match 93.7%; Score 521; DB 15; Length 128;  
Best Local Similarity 92.5%; Pred. No. 4.8e-30;  
Matches 98; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 QIVLSQSPAILASPGKVTMTCRASSSVNYMHYQOKPGSSPKPWISATSNLASGVPAR 60

DB 23 QIVLSQSPAILASPGKVTMTCRASSSVNYMHYQOKPGSSPKPWIVATSNLASGVPAR 82

OY 61 FSGSGSGTSTLTISRVEADATYCCQWSSNPPTFGGTMLEIR 106

DB 83 FSGSGSGTSTLTISRVEADATYCCQWSSNPPTFGGTMLEIR 128

#### RESULT 8

AAU30694

ID AAB30694 standard; Protein: 412 AA.  
 AC AAB30694;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE A fusion of anti-CD20 single chain antibody/streptavidin.  
 XX  
 KM Streptavidin; tumour cell; cancer; adenocarcinoma;  
 KM hematological malignancy; B9E9.  
 XX  
 OS Synthetic.  
 OS Streptomyces avidinii.  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Protein  
 FT /note= "VL chain"  
 FT 109..126  
 FT Peptide  
 FT /note= "linker"  
 FT 127..248  
 FT Protein  
 FT /note= "VH chain"  
 FT 249..253  
 FT Peptide  
 FT /note= "linker"  
 FT 254..412  
 FT Protein  
 FT /note= "streptavidin"  
 XX  
 FN WO20075333-A1.  
 XX  
 PD 14-DEC-2000.  
 XX  
 PF 05-JUN-2000; 2000WO-US15595.  
 XX  
 PR 07-JUN-1999; 99US-0137900.  
 PR 03-DEC-1999; 99US-0168976.  
 XX  
 PA (NEOR-) NEORX CORP.  
 XX  
 PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
 DR WPI: 2001-091213/10.  
 DR N-PSDB: AAC86563.  
 XX  
 PT New vector constructs for expressing genomic streptavidin fusion  
 PT proteins which are useful for targeting tumour cells associated with  
 PT cancer, e.g. adenocarcinomas -  
 XX  
 PS Example 2; Fig 11B; 100pp; English.  
 XX  
 CC The present sequence represents a fusion of an anti-CD20 single chain  
 CC antibody (B9E9) streptavidin. The fusion protein is expressed using  
 CC vectors of the invention. The specification describes vector constructs  
 CC for expressing streptavidin fusion proteins. The vector comprises a  
 CC nucleic acid encoding streptavidin or its functional variant operatively  
 CC linked to a promoter, and a cloning site for insertion of a second  
 CC nucleic acid sequence encoding a polypeptide to be fused with  
 CC streptavidin, interposed between the promoter and the first nucleic  
 CC acid sequence. Alternatively, the vector construct comprises a nucleic  
 CC acid, operatively linked to a promoter, encoding a polypeptide to be  
 CC fused with streptavidin, and a cloning site for insertion of a second  
 CC nucleic acid encoding at least 129 amino acids of streptavidin or its  
 CC functional variant. The fusion proteins are useful for targeting tumour  
 CC cells, particularly tumour cells associated with cancer,  
 CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
 CC is useful for expressing of streptavidin fusion proteins. In particular,  
 CC these are useful as tools for medical diagnostics and therapeutic  
 CC purposes, e.g. for detecting the presence or absence of, or treating, a  
 CC target site within a mammalian host.  
 CC  
 XX  
 S0 Sequence 412 AA;

Query Match 93.0%; Score 517; DB 22; Length 412;  
 Best Local Similarity 93.3%; Pred. NO. 2.8e-29;

Matches 98; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 OY 2 IVLSQSPAILLSASPEEKVTMTCRASSVYVMHWYQOKPSSPKPVLSATSNLASCVPARF 61  
 DB 2 IVLSQSPAILLSASPEEKVTMTCRASSVYVMHWYQOKPSSPKPVLSATSNLASCVPARF 61  
 OY 62 SGSGSGTSTSYLSIRVEADATYTCQOMSSNPFFGGTWTLEIK 106  
 DB 62 SGSGSGTSTSYLSIRVEADATYTCQOMSSNPFFGGTWTLEIK 106  
 RESUT 9  
 ID AAB30695 standard; Protein: 423 AA.  
 XX  
 AC AAB30695;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE A fusion of anti-CD20 single chain antibody/streptavidin.  
 XX  
 KM Streptavidin; tumour cell; cancer; adenocarcinoma;  
 KM hematological malignancy; B9E9.  
 XX  
 OS Synthetic.  
 OS Streptomyces avidinii.  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Protein  
 FT /note= "VH chain"  
 FT 127..150  
 FT Peptide  
 FT /note= "linker"  
 FT 151..259  
 FT Protein  
 FT /note= "VL chain"  
 FT 260..264  
 FT Peptide  
 FT /note= "linker"  
 FT 265..423  
 FT Protein  
 FT /note= "streptavidin"  
 XX  
 FN WO20075333-A1.  
 XX  
 PD 14-DEC-2000.  
 XX  
 PF 05-JUN-2000; 2000WO-US15595.  
 XX  
 PR 07-JUN-1999; 99US-0137900.  
 PR 03-DEC-1999; 99US-0168976.  
 XX  
 PA (NEOR-) NEORX CORP.  
 XX  
 PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
 DR WPI: 2001-091213/10.  
 DR N-PSDB: AAC86564.  
 XX  
 PT New vector constructs for expressing genomic streptavidin fusion  
 PT proteins which are useful for targeting tumour cells associated with  
 PT cancer, e.g. adenocarcinomas -  
 XX  
 PS Example 2; Fig 11C; 100pp; English.  
 XX  
 CC The present sequence represents a fusion of an anti-CD20 single chain  
 CC antibody (B9E9) streptavidin. The fusion protein is expressed using  
 CC vectors of the invention. The specification describes vector constructs  
 CC for expressing streptavidin fusion proteins. The vector comprises a  
 CC nucleic acid encoding streptavidin or its functional variant operatively  
 CC linked to a promoter, and a cloning site for insertion of a second  
 CC nucleic acid sequence encoding a polypeptide to be fused with  
 CC streptavidin, interposed between the promoter and the first nucleic  
 CC acid sequence. Alternatively, the vector construct comprises a nucleic  
 CC acid, operatively linked to a promoter, encoding a polypeptide to be  
 CC fused with streptavidin, and a cloning site for insertion of a second

[illegible][illegible]



```

FT Domain 110..118
FT /label=CDR3.
FT Domain 119..128
FT /label=FR4.
XX
XX WO8900999-A.
XX
XX 09-FEB-1989.
XX
XX 25-JUL-1988; 88WO-US02514.
XX
XX 24-JUL-1987; 87US-0077528.
XX
XX (ITGE-) INT GENETIC ENG INC.
XX
XX Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;
XX
XX WPI; 1989-061144/08.
XX
XX N-PSDB; AAN91147.
XX
XX Polynucleotide(s) encoding immunoglobulin molecules -
XX used for efficient produ. of chimeric human or non-human or
XX class switched antibodies.
XX
XX Disclosure; 7pp; English.
XX
XX Sequence carries the variable region of the chimeric immunoglobulin
XX sequence. The antibodies are useful in passive immunisation avoiding
XX negative immune reactions. They are also useful in assaying and in vitro
XX imaging.
XX
XX Sequence 128 AA:
XX
XX Query Match 92.8%; Score 516; DB 10; Length 128;
XX Best Local Similarity 92.5%; Pred. No. 1.1e-29;
XX Matches 98; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 QIVLSOSPAILLSASPEKXTMTCRASSSYNMHWYQOKFGSSPKPWISATSNLASGVPAR 60
DB 23 QIVLSOSPAILLSASPEKXTMTCRASSSYNMHWYQOKFGSSPKPWISATSNLASGVPAR 82
OY 61 FSGSGSGTSTSLTISRVEADATYTCQOWSSNPPTFGGTMLEIR 106
DB 83 FSGSGSGTSTSLTISRVEADATYTCQOWSSNPPTFGGTMLEIR 128

RESULT 12
AAM10589
ID AAM10589 standard; Protein: 128 AA.
XX
XX AAM10589;
XX
XX 21-OCT-1997 (first entry)
XX
XX 2H7 antibody light chain variable region.
XX
XX Immunoglobulin G; IgG; light chain; recombinant production;
XX antibody; passive immunisation; serum sickness; anaphylactic shock;
XX immunosassay; imaging; reagent; complement mediated lysis;
XX therapy; variable; region; complementarity determining; CDR;
XX framework.
XX
XX Mus spp.
XX
XX Key Location/Qualifiers
XX Peptide 1..22
XX /label= sig_peptide
XX Peptide 23..128
XX /label= mat_peptide
XX Region 23..46
XX /label= framework_1
XX Region 47..55
XX /label= CDR_1

```

```

FT Region 56..70
FT /label= framework_2
FT Region 71..77
FT /label= CDR_2
FT Region 78..109
FT /label= framework_3
FT Region 110..118
FT /label= CDR_3
FT Region 119..128
FT /label= framework_4
XX
XX US5595898-A.
XX
XX 21-JAN-1997.
XX
XX 01-NOV-1985; 85US-0793980.
XX
XX 29-MAR-1990; 90US-0501092.
XX
XX 01-NOV-1985; 85US-0793980.
XX
XX 27-OCT-1986; 86WO-US02269.
XX
XX 24-JUL-1987; 87US-0077528.
XX
XX 11-JAN-1988; 88US-0142039.
XX
XX 08-DEC-1992; 92US-0987555.
XX
XX 18-AUG-1994; 94US-0299085.
XX
XX (XOMA ) XOMA CORP.
XX
XX Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;
XX Wall R, Wilcox GL;
XX
XX WPI; 1997-107579/10.
XX
XX N-PSDB; AAT36317.
XX
XX Nucleic acid encoding immunoglobulin fragment - comprising
XX dicistronic transcription unit with pectate lyase signal sequences
XX
XX Example; Fig 22; 95pp; English.
XX
XX The present sequence is the light chain variable region of
XX the 2H7 antibody. The 2H7 cDNA was used in the preparation of
XX a novel polynucleotide molecule encoding an Ig fragment. The DNA
XX molecule comprises 2 DNA sequences encoding 2 pectate lyase
XX secretion signal sequences respectively linked to a DNA sequence
XX encoding an Ig Fd molecule or Ig light chain, operably linked to a
XX single prokaryotic promoter so as to form a dicistronic
XX transcription unit, provided that the Ig fragment can bind an
XX antigen and is produced and secreted by an E. coli host cell when
XX the nucleic acid molecule is expressed in the host cell.
XX The polynucleotide molecule is used for the production of
XX recombinant antibodies, which can be used for passive immunisation
XX without negative immune reactions (e.g. serum sickness and
XX anaphylactic shock), in labelled forms as immunoassay or imaging
XX reagents, in complement mediated lysis and for therapeutic
XX purposes when coupled to a toxin or other therapeutic agent.
XX
XX Sequence 128 AA:
XX
XX Query Match 92.8%; Score 516; DB 18; Length 128;
XX Best Local Similarity 92.5%; Pred. No. 1.1e-29;
XX Matches 98; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 QIVLSOSPAILLSASPEKXTMTCRASSSYNMHWYQOKFGSSPKPWISATSNLASGVPAR 60
DB 23 QIVLSOSPAILLSASPEKXTMTCRASSSYNMHWYQOKFGSSPKPWISATSNLASGVPAR 82
OY 61 FSGSGSGTSTSLTISRVEADATYTCQOWSSNPPTFGGTMLEIR 106
DB 83 FSGSGSGTSTSLTISRVEADATYTCQOWSSNPPTFGGTMLEIR 128

RESULT 13
AAM16344
ID AAM16344 standard; Protein: 128 AA.

```

```

XX AC AAW16344:
XX
XX 04-SEP-1997 (first entry)
XX DE 2H7 light chain variable sequence.
XX
XX Antibody engineering; heavy chain; light chain; chimaeric antibody;
XX passive immunisation; diagnosis; hybridoma; monoclonal antibody;
XX 2H7; B-cell antigen; Bp35.
XX
XX Mus sp.
XX
XX Key location/Qualifiers
XX FH 1..22
XX FT Peptide /label= Sig-peptide
XX FT 23..46
XX FT Region /label= FRL
XX FT /note= "framework region 1"
XX FT 47..55
XX FT Region /label= CDRI
XX FT /note= "complementarity determining region 1"
XX FT 56..70
XX FT Region /label= FR2
XX FT /note= "framework region 2"
XX FT 71..77
XX FT Region /label= CDR2
XX FT /note= "complementarity determining region 2"
XX FT 78..109
XX FT Region /label= FR3
XX FT /note= "framework region 3"
XX FT 110..118
XX FT Region /label= CDR3
XX FT /note= "complementarity determining region 3"
XX FT 119..128
XX FT /label= FR4
XX FT /note= "framework region 4"
XX
XX US5618920-A.
XX
XX 08-APR-1997.
XX
XX 01-NOV-1985; 85US-0793980.
XX
XX 29-MAR-1990; 90US-0501092.
XX PR 01-NOV-1985; 85US-0793980.
XX PR 27-OCT-1986; 86WO-US02269.
XX PR 24-JUL-1987; 87US-0077528.
XX PR 11-JAN-1988; 88US-0142039.
XX PR 17-APR-1992; 92US-0870404.
XX PR 29-APR-1994; 94US-0235225.
XX
XX (XOMA ) XOMA CORP.
XX
XX Better M, Horwitz AH, Lei S, Liu AY, Robinson RR,
XX Wall R, Wilcox GL;
XX
XX WPI; 1997-225473/20.
XX DR N-PSDB; AAT70869.
XX
XX Secretable immunoglobulin heavy and light chain fragments - capable
XX of assembling into chimeric antibodies, useful for e.g. passive
XX immunisation, diagnosis, etc
XX
XX Example 4; Fig 22; 96pp; English.
XX
XX A polypeptide (AAW16344) comprises the light chain variable region of
XX mouse monoclonal antibody 2H7, which recognises human B-cell surface
XX antigen Bp35. It is encoded by a cDNA clone (AAT70869) obtd. from a
XX 2H7 cDNA library by PCR amplification. The 2H7 heavy chain variable
XX sequence (AAW16343) is also provided. A novel human-mouse chimaeric
XX antibody with specificity for the human B-cell antigen has been
XX constructed.

```

```

XX SQ Sequence 128 AA:
XX
XX Query Match 92.8%; Score 516; DB 18; Length 128;
XX Best Local Similarly 92.5%; Pred. No: 1; Le-29;
XX Matches 98; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 QIVSQSPALISASPGKVTMTCPASSSVNMMHYQOKPGSSPKPVTSATSNLASGVPAR 60
XX DB 23 QIVSQSPALISASPGKVTMTCPASSSVNMMHYQOKPGSSPKPVTSATSNLASGVPAR 82
XX
XX 61 FSGSGSGTSTSLTISRVEAEDATYYCOQSSNPTEFGGTMLEIR 106
XX DB 83 FSGSGSGTSTSLTISRVEAEDATYYCOQSSNPTEFGGTMLEIR 128
XX
XX RESULT 14
XX ID AAW10243 standard; Protein; 128 AA.
XX AC AAW10243:
XX
XX 06-AUG-1997 (first entry)
XX
XX Light chain variable sequence of 2H7.
XX
XX Pectate lyase; signal sequence; Gram-negative bacterium; immunoglobulin;
XX protein production; human; constant region; passive immunisation; toxin;
XX serum sickness; anaphylaxis; sweetener; thrombatin; cytoplasm; periplasm;
XX antibody; Ig; light-chain; hepatitis; mouse; lung carcinoma; cancer;
XX myeloma cell.
XX
XX Mus musculus.
XX
XX Key location/Qualifiers
XX FH 1..22
XX FT Peptide /note= "leader peptide"
XX FT 23..46
XX FT Region /note= "framework region 1"
XX FT 47..55
XX FT Region /note= "complementarity determining region 1"
XX FT 56..70
XX FT Region /note= "framework region 2"
XX FT 71..77
XX FT Region /note= "complementarity determining region 2"
XX FT 78..109
XX FT Region /note= "framework region 3"
XX FT 110..118
XX FT Region /note= "complementarity determining region 3"
XX FT 119..128
XX FT /note= "framework region 4"
XX
XX US5576195-A.
XX
XX 19-NOV-1996.
XX
XX 01-NOV-1985; 85US-0793980.
XX
XX 29-MAR-1990; 90US-0501092.
XX PR 01-NOV-1985; 85US-0793980.
XX PR 27-OCT-1986; 86WO-US02269.
XX PR 24-JUL-1987; 87US-0077528.
XX PR 11-JAN-1988; 88US-0142039.
XX PR 08-DEC-1992; 92US-0987555.
XX PR 22-FEB-1993; 93US-0020671.
XX PR 09-DEC-1994; 94US-0357234.
XX
XX (XOMA ) XOMA CORP.
XX
XX Better M, Lei S, Robinson RR, Wilcox GL;
XX WPI; 1997-011254/01.
XX DR N-PSDB; AAT51043.

```

XX Improved prodn. of protein in Gram -ve bacteria using: signal  
 PT sequence - from pectate lyase to ensure transport of protein from  
 PT the cytoplasm, esp. for prodn. of antibodies  
 XX  
 PS Example 4; Fig 22; 86pp; English.

XX This sequence represents the light chain variable region of the 2H7  
 CC mouse monoclonal antibody. The 2H7 antibody recognises the human B-cell  
 CC surface antigen Bp35, which plays a role in B-cell activation. This  
 CC sequence was used in a human-mouse chimeric antibody with human B-cell  
 CC antigen specificity, that was produced using the method of the invention.  
 CC The method of the invention is for the production of a protein in a  
 CC Gram-negative bacterium. The method improves on current techniques, by  
 CC using a vector including DNA encoding the pectate lyase signal sequence  
 CC (see AAT51034), attached to the sequence encoding the protein for  
 CC production. The method is especially used to make immunoglobulins (Ig),  
 CC particularly those with a human constant region, suitable for passive  
 CC immunisation (without risk of serum sickness or anaphylaxis) or for in  
 CC vivo/in vitro diagnosis and imaging. The Ig may also be used  
 CC therapeutically, optionally coupled to toxins, etc. Alternatively the  
 CC protein to be produced is the sweetener thaumatin. The presence of this  
 CC signal sequence means that the protein is exported from the cytoplasm  
 CC and can be recovered from the culture medium or periplasm, in active and  
 CC correctly folded form. The method allows the class of any antibody to be  
 CC switched, e.g. most human-human Ig are of M class, easily reduced and  
 CC aggregated, and these can now be changed to G, A or E classes.

XX Sequence 128 AA:

Query Match 92.8%; Score 516; DB 18; Length 128;  
 Best Local Similarity 92.5%; Pred. No. 1.le-29;  
 Matches 98; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 QIVLSOSPAILASPEGEKVTMTCRASSVYMHMWYQOKFGSSPKPWISATSNLASCVPAR 60  
 DB 23 QIVLSOSPAILASPEKVTMTCRASSVSYMHMWYQOKFGSSPKPWIVAPSNLASCVPAR 82  
 OY 61 FSGSGGTSYSLTISRVEADAATYYCQOWSSNPFFGCGTMEIR 106  
 DB 83 FSGSGGTSYSLTISRVEADAATYYCQOWSNPFFGAGTKLEIK 128

RESULT 15  
 AAM47521  
 ID AAM47521 standard: Protein; 128 AA.

XX AAM47521;

DT 05-JUN-1998 (first entry)

XX Mouse 2H7 antibody light chain variable region.

XX Mouse; murine; light chain; variable region;  
 KW Immunoglobulin fragment production; Ig fragment production;  
 KM monoclonal antibody 2H7; human B-cell surface antigen.

XX Mus sp.

XX Key Location/Qualifiers  
 FT 1..22  
 FT /label= sig\_peptide  
 FT 23..128  
 FT /label= mat\_peptide

XX US5698417-A.

XX 16-DEC-1997.

XX 06-JUN-1995; 95US-0466203.

XX 29-MAR-1990; 90US-0501092.  
 PR 01-NOV-1985; 85US-0793980.

PR 27-OCT-1986; 86WO-US02269.  
 PR 24-JUL-1987; 87US-0077528.  
 PR 11-JAN-1988; 88US-0142039.  
 PR 08-DEC-1992; 92US-0987555.  
 PR 18-AUG-1994; 94US-0299085.  
 PR 25-MAY-1995; 95US-0450731.  
 PR 06-JUN-1995; 95US-0466203.

PA (XOMA ) XOMA CORP.

PI Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;  
 PI Wall R, Wilcox GL;

DR WPI: 1998-051487/05.

DR N-PSDB: AAV18594.

PT Production of recombinant immunoglobulin fragment - comprising Fd  
 PT molecule and light chain

PS Example IV; Fig 22; 98pp; English.

XX The present sequence was used in the development of a novel method  
 CC for the production of an immunoglobulin (Ig) fragment capable of  
 CC binding an antigen. The method comprises culturing an E. coli host  
 CC that has been transformed with a nucleic acid molecule encoding the  
 CC Ig fragment, under conditions so that the Ig fragment is produced  
 CC and secreted. The nucleic acid molecule comprises DNA sequences  
 CC encoding: (a) pectate lyase secretion signal sequence operably  
 CC linked to a DNA sequence encoding at least the variable region of  
 CC an Ig Fd molecule; and (b) pectate lyase secretion signal sequence  
 CC operably linked to a DNA sequence encoding at least the variable  
 CC region of an Ig light chain, where (a) and (b) are operably linked  
 CC to a single prokaryotic promoter to form a dicistronic  
 CC transcription unit. The method is used to produce chimeric Fab  
 CC molecules, e.g. derived from murine monoclonal antibody 2H7 raised  
 CC against human B-cell surface antigen. The invention provides a  
 CC novel approach for producing genetically engineered antibodies of  
 CC desired variable region specificity and constant region  
 CC properties. The cloned Ig gene products can be produced by  
 CC expression in genetically engineered organisms. The application of  
 CC chemical gene synthesis, recombinant DNA cloning and production of  
 CC specific Ig chains in various organisms provides an effective  
 CC solution for the efficient large scale production of human  
 CC monoclonal antibodies. The invention also provides a solution to  
 CC the problem of class switching antibody molecules.

XX Sequence 128 AA:

Query Match 92.8%; Score 516; DB 19; Length 128;  
 Best Local Similarity 92.5%; Pred. No. 1.le-29;  
 Matches 98; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 QIVLSOSPAILASPEGEKVTMTCRASSVYMHMWYQOKFGSSPKPWISATSNLASCVPAR 60  
 DB 23 QIVLSOSPAILASPEKVTMTCRASSVSYMHMWYQOKFGSSPKPWIVAPSNLASCVPAR 82  
 OY 61 FSGSGGTSYSLTISRVEADAATYYCQOWSSNPFFGCGTMEIR 106  
 DB 83 FSGSGGTSYSLTISRVEADAATYYCQOWSNPFFGAGTKLEIK 128

Search completed: November 27, 2002, 07:17:57  
 Job time : 62.5251 secs

THIS PAGE BLANK (18970)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:15:05 ; Search time 29.0579 seconds

(without alignments)  
350.688 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 556  
Sequence: 1 QIVLSQSPALISAPGEKVT.....CQKMSNPPTFGGCTMLEIR 106Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 100%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_73:.\*  
2: PIR1:.\*  
3: PIR2:.\*  
4: PIR3:.\*  
5: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	523	94.1	106	2	PI0082
2	492	88.5	107	2	A42848
3	480	86.3	107	2	PC4405
4	474	85.3	107	2	A30362
5	469	84.4	107	2	S26338
6	465.5	83.9	108	2	B30562
7	465.5	83.7	108	2	G30560
8	462	83.1	140	2	S38720
9	459	82.6	106	2	PL0013
10	458	82.4	104	2	PS0071
11	458	82.4	107	2	B49049
12	458	82.4	130	1	PD0011
13	457	82.2	103	2	JL0079
14	454	81.7	100	2	S29591
15	454	81.7	100	2	S29580
16	454	81.7	107	2	PT0406
17	452	81.3	97	2	S26341
18	446	80.2	106	2	B54378
19	445	80.0	235	2	S25058
20	444	79.9	107	2	S11118
21	442.5	79.6	104	2	TC6076
22	442	79.5	107	2	S11119
23	440	79.1	130	2	A32513
24	437	78.6	97	2	PH1084
25	434	78.1	107	2	S11121
26	433	77.9	107	2	S11121
27	433	77.9	130	2	B32456
28	433	77.9	130	2	S04573
29	430	77.3	99	2	S29585

30	429	77.2	106	2	G27887	Ig kappa chain V r
31	427	76.8	97	2	PH1085	Ig light chain V r
32	426	76.6	94	2	S26340	Ig light chain V r
33	426	76.6	99	2	D38601	Ig kappa chain V r
34	425	76.4	109	2	PT0405	Ig light chain V r
35	423	76.1	95	2	D33730	Ig kappa chain V r
36	423	76.1	107	2	S11112	Ig kappa chain V r
37	423	76.1	107	2	S11113	Ig kappa chain V r
38	423	76.1	108	2	S29581	Ig kappa chain V r
39	421	75.7	108	2	PL0278	Ig kappa chain V r
40	420	75.5	109	2	PT0404	Ig light chain V r
41	420	75.5	132	2	S05268	Ig kappa chain pre
42	419	75.4	106	2	PS0070	Ig kappa chain V r
43	418	75.2	107	1	KYMSY4	Ig kappa chain V r
44	417	75.0	108	2	PL0277	Ig kappa chain V r
45	417	75.0	108	2	PL0277	Ig kappa chain V r

## ALIGNMENTS

## RESULT 1

PL0082  
Ig kappa chain V region (2D3) - mouseC:Species: Mus musculus (house mouse)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 21-Jan-2000

C:Accession: PL0082

R:Meek, K.; Hasemann, C.; Pollok, B.; Alken, S.S.; Bratt, M.; Slouli, M.; Urbain, J.;

J. Exp. Med. 169, 519-533, 1989

A:Title: Structural characterization of antidiabetic antibodies; evidence that Ab2s

A:Reference number: PL0080; MUID:89094248; PMID:2492056

A:Accession: PL0082

A:Molecule type: mRNA

A:Residues: 1-106 &lt;ME&gt;

A:Experimental source: Strain BALB/c

A:Note: the sequence shown here is from the V kappa region of an antidiabetic monoclonal

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin

F:16-89/Domain: Immunoglobulin homology &lt;IM&gt;

Query Match 94.1%; Score 523; DB 2; Length 106;  
Best Local Similarity 94.3%; Pred. No. 9.2e-38;

Matches 100; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 OIVLSQSPALISAPGEKVTTCRASSSVYHWQKPGSSPKPMISATSLASGVAPR 60

DB 1 OIVLSQSPALISAPGEKVTTCRASSSVYHWQKPGSSPKPMISATSLASGVAPR 60

OY 61 FSGSGSGTSTLTISRVEAEDATATTCQWSSNPPTFGGCTMLEIR 106

DB 61 FSGSGSGTSTLTISRVEAEDATATTCQWSSNPPTFGGCTMLEIR 106

## RESULT 2

A42848  
Ig light chain V region - mouse (fragment)C:Species: Mus musculus (house mouse)  
C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000

C:Accession: A42848; S33902

R:Fell, H.P.; Gayle, M.A.; Yelton, D.; Lipsitch, L.; Schieven, G.L.; Marken, J.S.; Aru

J. Biol. Chem. 267, 15552-15558, 1992

A:Title: Chimeric Ig anti-tumor antibody. Genomic construction, expression, and chara

A:Reference number: A42848; MUID:92348410; PMID:1639794

A:Accession: A42848

A:Status: Preliminary

A:Molecule type: DNA

A:Cross-references: EMBL:M90690  
A:Note: sequence extracted from NCBI backbone (NCBI:109958, NCBI:109959)  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: Immunoglobulin  
F:16-89/Domain: Immunoglobulin homology <IM>



```

A:Title: Variable region cDNA sequences and antigen binding specificity of mouse monoclonal
A:Reference number: A30560; MUID:89110062; PMID:2464028
A:Accession: G30560
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-108 <MAT>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:16-89/Domain: Immunoglobulin homology <IMM>

Query Match      83.9%   Score 466.5;   DB 2;   Length 108;
Best Local Similarity 85.0%;   Pred. No. 5.8e-33;
Matches 91;   Conservative 5;   Mismatches 10;   Indels 1;   Gaps 1;

OY 1 QIVLSQSPAILTSASGEKVMTCTCRASSSVNYMHVYQOKPGSSPKPMISATSNLASGVPAR 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QIVLTQSPALMSASPGKEVMTCTCASSSVSYMHVYQOKSGSSPKTWIDYTKLASGVPAR 60

OY 61 FSGSGSGTYSLTISIRVAEADATYYCOQWSSNP-TRGGGTMLEIR 106
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 FSGSGSGTYSLTISIRVAEADATYYCOQWTRNPETFGSGTKLEIR 107

RESULT 8
S38720
Ig light chain V region - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
C:Accession: S38720
R:Climanis, A.Y.
submitted to the EMBL Data Library, November 1993
A:Reference number: S38713
A:Accession: S38720
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-108 <CTIM>
A:Cross-references: EMBL:X76023; NID:9416104; PIDN:CAAS3610.1; PID:g1334265
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: Immunoglobulin
F:16-89/Domain: Immunoglobulin homology <IMM>

Query Match      83.7%   Score 465.5;   DB 2;   Length 108;
Best Local Similarity 82.2%;   Pred. No. 7e-33;
Matches 88;   Conservative 10;   Mismatches 8;   Indels 1;   Gaps 1;

OY 1 QIVLSQSPAILTSASGEKVMTCTCRASSSVNYMHVYQOKPGSSPKPMISATSNLASGVPAR 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EIVLTQSPALMSASPGKEVMTCTCASSSVSYMHVYQOKPRSSPKRWILTISNLASGVPLR 60

OY 61 FSGSGSGTYSLTISIRVAEADATYYCOQWSSNP-TRGGGTMLEIR 106
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 FSGSGSGTYSLTISIRVAEADATYYCOQWSSNPPLTRTGAGTKQLK 107

RESULT 9
PL0013
Ig kappa chain precursor V region (4C11) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Jan-2000
C:Accession: PL0013
R:Cheng, H.L.; Sood, A.K.; Ward, R.E.; Kleber-Emmons, T.; Kohler, H.
Mol. Immunol. 25, 33-40, 1988
A:Title: Structural basis of stimulatory anti-idiotypic antibodies.
A:Reference number: PL0011; MUID:88142863; PMID:3125424
A:Accession: PL0013
A:Molecule type: mRNA
A:Residues: 1-140 <CHD>
A:Experimental source: cell line 4C11
A:Comment: This protein is an anti-idiotypic antibody that induces an anti-phosphorylcholine
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-32/Domain: signal sequence #status predicted <SIG>

```

```

F:23-129/Product: Ig heavy chain V region (4C11) #status predicted <MAT>
F:38-111/Domain: Immunoglobulin homology <IMM>
F:46-55/Region: complementarity-determining 1
F:71-77/Region: complementarity-determining 2
F:110-118/Region: complementarity-determining 3
F:130-140/Domain: constant region (fragment) #status predicted <CON>

Query Match      83.1%   Score 462;   DB 2;   Length 140;
Best Local Similarity 83.0%   Pred. No. 1.8e-32;
Matches 88;   Conservative 10;   Mismatches 32;   Indels 0;   Gaps 0;

OY      1  QIVLSQSPALLSASGPEVYTMCRASSSVNVMHWYQOKFGSSPKPMISATSNLASGVPAR 60
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      23  QIVLQSPALMSASGGEVYITCSASSSVSYMHMFQOKPDPISPKIMITYSTNLSAGVPR 82
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

OY      61  FSGSGSGTYSYLTISRVEADATATYCCOOWSSNPPTFCGGMLEIR 106
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      83  FSGSGSGTYSYLTISRMEADATATYCCQRRSSYPPTFCGGKLEIK 128
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
PS0071
Ig kappa chain V region (38C13_V8) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jan-2000
C:Accession: PS0071
R:Levy, S.; Campbell, M.J.; Levy, R.
J. Exp. Med. 170, 1-13, 1989
A:Title: Functional immunoglobulin light chain genes are replaced by ongoing rearrang
A:Reference number: A92781; MUID:89310348; PMID:2501443
A:Accession: PS0071
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-106 <LEV>
C:Superfamily: Immunoglobulin V region: Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:16-89/Domain: Immunoglobulin homology <IMM>

Query Match      82.6%   Score 459;   DB 2;   Length 106;
Best Local Similarity 83.0%   Pred. No. 2.4e-32;
Matches 88;   Conservative 7;   Mismatches 11;   Indels 0;   Gaps 0;

OY      1  QIVLSQSPALLSASGCEVYTMCRASSSVNVMHWYQOKFGSSPKPMISATSNLASGVPAR 60
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      1  QIVLQSPALMSASGGEVYTMCRASSSVSYMHYQOKGSGSPKMIYDTSKLASGVPAR 60
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

OY      61  FSGSGSGTYSYLTISRVEADATATYCCOOWSSNPPTFCGGMLEIR 106
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      61  FSGSGSGTYSYLTISRMEADATATYCCOOWSSNPPTFCGAPKLEIK 106
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
BA9049
Ig kappa chain V region (anti-idiotypic) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
C:Accession: BA9049
R:Armandola, E.A.; Martini, S.M.; Zwickl, M.; Hardman, N.; Ferrone, S.
Eur. J. Immunol. 22, 2893-2899, 1992
A:Title: Molecular analysis of anti-idiotypic monoclonal antibodies in the HLA-DR anti
A:Reference number: A49049; MUID:93049629; PMID:1425914
A:Accession: BA9049
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-104 <ARM>
A:Experimental source: BALB/c
A>Note: sequence extracted from NCBI backbone (NCBIN:118298, NCBIPI:118299)
C:Superfamily: Immunoglobulin V region: Immunoglobulin homology
C:Keywords: immunoglobulin
F:16-89/Domain: Immunoglobulin homology <IMM>

Query Match      82.4%   Score 458;   DB 2;   Length 104;
Best Local Similarity 84.6%   Pred. No. 2.9e-32;

```





F:16-89/Domain: Immunoglobulin homology &lt;IMM&gt;

Query Match 81.7%; Score 454; DB 2; Length 100;

Best Local Similarity 87.0%; Pred. No. 6.1e-32;

Matches 87; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```
QY 1 QIVLSQSPAILASAPGEKVTMTCRASSSVNYHWYQOKFGSSPKPWISATSNLAGVPPAR 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 QIVLTQSPAIMSAPGEKVTMTCSASSSVYHWYQOKSGTSPKRWIYDTSKLAGVPPAR 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 FSGSGSGTSTSLTISRVEAEDAATYYCOQWSSNPPTFGGG 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 FSGSGSGTSTSLTISRVEAEDAATYYCOQWSSNPPTFGGG 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Search completed: November 27, 2002, 07:24:16  
Job time : 30.0579 secs

THIS PAGE BLANK (USPTO)

DD I QIVLQSPALMSASFGUAVIMICLSASSSVSIMHWIQDAS

2000

61 **ESGSGSGT**SYSLT**SRVEA**EDATYYCOWS**SNPPTGGGTM**.ETR 106

QY 61 FSGSGGTSYSLTISRVEADATYCCQWSSNPPTFGGGMLEIR 106

Db 61 FSGSGSATSYSLTITTSMQADAAATYYCQOWSSNPLTFGAGTKLELK 106

## RESULT 2

KV6L\_MOUSE

STANDARD:

PRT: 107 AA.

AC P04943;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig kappa chain V-VI region NQ6-8.3.1.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBITaxID=10090;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83271467; PubMed=6877353;  
 RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
 RT "mRNA sequences define an unusually restricted IgG response to 2-phenylloxazalone and its early diversification."  
 RL Nature 304:320-324(1983).  
 CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
 -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).  
 -----

CC EMBL: K00740; AAA38685.1; -  
 DR HSSP: P01679; 2FBJ.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_1.  
 DR SMART: SM00406; IGV\_1.  
 KW Immunoglobulin V region; Hybridoma.  
 FT DOMAIN 1 23 FRAMEWORK-1.  
 FT DOMAIN 24 33 COMPLEMENTARITY-DETERMINING-1.  
 FT DOMAIN 34 48 FRAMEWORK-2.  
 FT DOMAIN 49 55 COMPLEMENTARITY-DETERMINING-2.  
 FT DOMAIN 56 87 FRAMEWORK-3.  
 FT DOMAIN 88 96 COMPLEMENTARITY-DETERMINING-3.  
 FT DOMAIN 97 106 FRAMEWORK-4.  
 FT DISULFID 23 87 BY SIMILARITY.  
 FT NON\_TER 107 107  
 SQ SEQUENCE 107 AA; 11572 MW; 6F694824ECF0C8E6 CRC64;

Query Match 80.2%; Score: 446; DB 1; Length 107;  
 Best Local Similarity 79.2%; Pred. No. 1.2e-35;  
 Matches 84; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

OY 1 QIVLSOSPALLSAPGKVTMTCRASSVYVMHYQKPPSSPKRWISATSNLASGVPAR 60  
 DB 1 QIVLTOSPALMSAPGKVTMTCSASSSVYMHYQKSGTSPKRWIYDTSKLASGVPAR 60  
 OY 61 FSGSGSATSYSLTITTSRVEADAATYYCQOWSSNPLTFGAGTKLEIR 106  
 DB 61 FSGSGSATSYSLTITTSMQADAAATYYCQOWSSNPLTFGAGTKLELK 106

## RESULT 3

KV6L\_MOUSE

STANDARD:

PRT: 107 AA.

AC P04942;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig kappa chain V-VI region NQ5-61.1.2.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBITaxID=10090;  
 RP SEQUENCE FROM N.A.

RX MEDLINE=83271467; PubMed=6877353;  
 RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
 RT "mRNA sequences define an unusually restricted IgG response to 2-phenylloxazalone and its early diversification."  
 RL Nature 304:320-324(1983).  
 CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
 -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

CC EMBL: K00739; AAA38684.1; -  
 DR HSSP: P01679; 2FBJ.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_1.  
 DR SMART: SM00406; IGV\_1.  
 KW Immunoglobulin V region; Hybridoma.  
 FT DOMAIN 1 23 FRAMEWORK-1.  
 FT DOMAIN 24 33 COMPLEMENTARITY-DETERMINING-1.  
 FT DOMAIN 34 48 FRAMEWORK-2.  
 FT DOMAIN 49 55 COMPLEMENTARITY-DETERMINING-2.  
 FT DOMAIN 56 87 FRAMEWORK-3.  
 FT DOMAIN 88 96 COMPLEMENTARITY-DETERMINING-3.  
 FT DISULFID 23 87 BY SIMILARITY.  
 FT NON\_TER 107 107  
 SQ SEQUENCE 107 AA; 11605 MW; CA6C4264ECCB550 CRC64;

Query Match 80.0%; Score 445; DB 1; Length 107;  
 Best Local Similarity 79.2%; Pred. No. 1.5e-35;  
 Matches 84; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

OY 1 QIVLSOSPALLSAPGKVTMTCRASSVYVMHYQKPPSSPKRWISATSNLASGVPAR 60  
 DB 1 QIVLTOSPALMSAPGKVTMTCSASSSVYMHYQKSGTSPKRWIYDTSKLASGVPAR 60  
 OY 61 FSGSGSATSYSLTITTSRVEADAATYYCQOWSSNPLTFGAGTKLEIR 106  
 DB 61 FSGSGSATSYSLTITTSMQADAAATYYCQOWSSNPLTFGAGTKLELK 106

## RESULT 4

KV6G\_MOUSE

STANDARD:

PRT: 107 AA.

AC P04941;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig kappa chain V-VI region NQ2-48.2.2.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBITaxID=10090;  
 RP SEQUENCE FROM N.A.

RX MEDLINE=83271467; PubMed=6877353;  
 RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
 RT "mRNA sequences define an unusually restricted IgG response to 2-phenylloxazalone and its early diversification."  
 RL Nature 304:320-324(1983).  
 CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
 -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: K00737: AAA38682.1; -  
DR HSSP: P01679: 2FBJ.  
DR InterPro: IPR003006: Ig\_MHC.  
DR Pfam: PF00047: Ig\_V.  
DR SMART: SM00406: Igy\_1.  
KW Immunoglobulin V region; Hybridoma.  
FT DOMAIN 1 23  
FT DOMAIN 24 33  
FT DOMAIN 34 48  
FT DOMAIN 49 55  
FT DOMAIN 56 87  
FT DOMAIN 88 96  
FT DOMAIN 97 106  
FT DISULFID 23 87  
FT NON\_TER 107 107  
SQ SEQUENCE 107 AA: 11556 MW: 72488DA9EF354934 CRC64;

Query Match  
Best Local Similarity 79.5%; Score 442; DB 1; Length 107;  
Matches 83; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 QIVLSQSPALISAPGKVTMTCRASSSVYMHVYQOKPSSPKPMISATSNLASGVPAR 60  
1 QILTLSPALMSASPGCKVTMTCSASSSVYMHVYQOKSGSPRMVYDTSKLASGVPAR 60  
DB 61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106  
61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106  
DB 61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106  
61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106

RESULT 5  
KV6K\_MOUSE STANDARD; PRT; 107 AA.  
AC P04944;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-VI region N05-78.2.6.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83271467; PubMed=6877353;  
RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
RT "mRNA sequences define an unusually restricted Igg response to 2-  
phenylloxazalone and its early diversification.";  
RL Nature 304:320-324(1983).  
CC - FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: K00744: AAA38689.1; -  
DR HSSP: P01679: 2FBJ.  
DR InterPro: IPR003006: Ig\_MHC.  
DR Pfam: PF00047: Ig\_V.  
DR SMART: SM00406: Igy\_1.

KW Immunoglobulin V region; Hybridoma.  
FT DOMAIN 1 23  
FT DOMAIN 24 33  
FT DOMAIN 34 48  
FT DOMAIN 49 55  
FT DOMAIN 56 87  
FT DOMAIN 88 96  
FT DOMAIN 97 106  
FT DISULFID 23 87  
FT NON\_TER 107 107  
SQ SEQUENCE 107 AA: 11611 MW: A38290781F3C30D3 CRC64;

Query Match  
Best Local Similarity 79.3%; Score 441; DB 1; Length 107;  
Matches 83; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 QIVLSQSPALISAPGKVTMTCRASSSVYMHVYQOKPSSPKPMISATSNLASGVPAR 60  
1 QILTLSPALMSASPGCKVTMTCSASSSVYMHVYQOKSGSPRMVYDTSKLASGVPAR 60  
DB 61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106  
61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106  
DB 61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106  
61 FSSGSGSTSLTISRVEADDAITYCOQWSSNPPTGGGTMLEIR 106

RESULT 6  
KV6K\_MOUSE STANDARD; PRT; 108 AA.  
AC P04945;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-VI region N02-6.1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83271467; PubMed=6877353;  
RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
RT "mRNA sequences define an unusually restricted Igg response to 2-  
phenylloxazalone and its early diversification.";  
RL Nature 304:320-324(1983).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: K00746: AAA38691.1; -  
DR HSSP: P01679: 2FBJ.  
DR InterPro: IPR003006: Ig\_MHC.  
DR Pfam: PF00047: Ig\_V.  
DR SMART: SM00406: Igy\_1.  
KW Immunoglobulin V region; Hybridoma.  
FT DOMAIN 1 23  
FT DOMAIN 24 33  
FT DOMAIN 34 48  
FT DOMAIN 49 55  
FT DOMAIN 56 87  
FT DOMAIN 88 98  
FT DOMAIN 99 108  
FT DISULFID 23 87  
FT NON\_TER 108 108  
SQ SEQUENCE 108 AA: 11713 MW: DABF235CD9680AC6 CRC64;

Query Match  
Best Local Similarity 78.2%; Score 435; DB 1; Length 108;  
Matches 83; Conservative 10; Mismatches 13; Indels 0; Gaps 0;



```

FT DOMAIN 49 55 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 56 87 FRAMEWORK-3.
FT DOMAIN 88 96 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 97 106 FRAMEWORK-4.
FT DISULFID 23 87 BY SIMILARITY.
FT NON_TER 107 107
SQ SEQUENCE 107 AA: 11584 MW: 36E6D02A5EC34D7 CRC64:

Query Match 74.3%; Score 413; DB 1; Length 107;
Best Local Similarity 75.5%; Pred. No. 1.6e-32;
Matches 80; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

Oy 1 QVLSQSPALISASGEKVTMTCRASSSVNYMHVYQOKRGSPPKPMISATSNLASGVPAR 60
Db 1 EIVLQSPALITASLGQKVTITCSASSSVSYMHVYQOKSGTSPKPMIYISKLASGVPAR 60

Oy 61 FSGSGSGTYSILTISRVEADATYYCOQWSSNPFTGGTMLEIR 106
Db 61 FSGSGSGTYSILTISRVEADATYYCOQWSSNPFTGGTMLEIR 106

RESULT 10
KV6C_MOUSE STANDARD: PRT; 107 AA.
ID KV6C_MOUSE
AC P01677;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-VI region TEPC 601/TEPC 191.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE (TEPC 601).
RX MEDLINE=79082830; PubMed=103573;
RA Rao D.N., Rudkoff S., Potter M.;
RT "k Chain variable regions from three galactan binding myeloma
RT proteins";
RL Biochemistry 17:555-5559(1978).
RN [2]
RP SEQUENCE (TEPC 191).
RX MEDLINE=81054757; PubMed=6776525;
RA Rudkoff S., Rao D.N., Glaudemans C.P.J., Potter M.;
RT "kappa Chain joining segments and structural diversity of antibody
RT combining sites";
RL Proc. Natl. Acad. Sci. U.S.A. 77:4270-4274(1980).
CC -1- MISCELLANEOUS: THE TWO SEQUENCES ARE IDENTICAL.
CC -1- MISCELLANEOUS: THESE CHAINS WERE ISOLATED FROM MYELOMA PROTEINS
CC THAT BIND GALACTAN.
PIR: A01941; KVMX4.
DR HSSP: P01679; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV_1.
KM Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 33 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 34 48 FRAMEWORK-2.
FT DOMAIN 49 55 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 56 87 FRAMEWORK-3.
FT DOMAIN 88 96 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 23 87 FRAMEWORK-4.
FT NON_TER 107 107 BY SIMILARITY.
SQ SEQUENCE 107 AA: 11568 MW: 203CDV52A5EC34D7 CRC64:

Query Match 73.9%; Score 411; DB 1; Length 107;
Best Local Similarity 74.5%; Pred. No. 2.5e-32;
Matches 79; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

Oy 1 QVLSQSPALISASGEKVTMTCRASSSVNYMHVYQOKRGSPPKPMISATSNLASGVPAR 60
```

```

Db 1 EIVLQSPALITASLGQKVTITCSASSSVSYMHVYQOKSGTSPKPMIYISKLASGVPAR 60

Oy 61 FSGSGSGTYSILTISRVEADATYYCOQWSSNPFTGGTMLEIR 106
Db 61 FSGSGSGTYSILTISRVEADATYYCOQWSSNPFTGGTMLEIR 106

RESULT 11
KV4A_MOUSE STANDARD: PRT; 129 AA.
ID KV4A_MOUSE
AC P01680;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region S107B precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82113500; PubMed=6799208;
RA Kwan S.-P., Max E.E., Seidman J.G., Leder P., Scharif M.D.;
RT "Two kappa immunoglobulin genes are expressed in the myeloma S107.";
RL Cell 26:57-66(1981).
CC -1- MISCELLANEOUS: THIS PROTEIN, IN WHICH THERE IS A DELETION OF TWO
CC AMINO ACIDS AT THE V-J RECOMBINATION SITE (AFTER POSITION 118), IS
CC SYNTHESIZED BUT NOT SECRETED IN CELLS THAT EXPRESS AND SECRETE THE
CC NORMAL KAPPA CHAIN S107.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: J00577; AAA8780.1; -
DR EMBL: V00780; CAA24157.1; -
DR PIR: A01943; KVM57B.
DR HSSP: P01679; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV_1.
KM Immunoglobulin V region; Signal.
FT SIGNAL 1 22
FT CHAIN 23 129 IG KAPPA CHAIN V-IV REGION S107B.
FT DOMAIN 24 45 FRAMEWORK-1.
FT DOMAIN 46 57 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 58 72 FRAMEWORK-2.
FT DOMAIN 73 79 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 80 111 FRAMEWORK-3.
FT DOMAIN 112 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 128 FRAMEWORK-4.
FT DISULFID 45 111 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA: 13833 MW: E4B73072DC6BE4 CRC64:

Query Match 72.8%; Score 405; DB 1; Length 129;
Best Local Similarity 77.4%; Pred. No. 1.1e-31;
Matches 82; Conservative 8; Mismatches 12; Indels 4; Gaps 2;

Oy 3 VLSQSPALISASGEKVTMTCRASSSV--NYMHVYQOKRGSPPKPMISATSNLASGVPAR 60
Db 25 VLTQSPALITASLGQKVTITCSASSSVSYLHMVYQOKSGASPKPIYIHTSNLASGVPAR 84

Oy 61 FSGSGSGTYSILTISRVEADATYYCOQWSSNPFTGGTMLEIR 106
Db 85 FSGSGSGTYSILTISRVEADATYYCOQWSSNPFTGGTMLEIR 128
```

RESULT 12  
KV3E\_MOUSE  
ID KV3E\_MOUSE STANDARD: PRT: 107 AA.  
AC P01679;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 19 kappa chain V-VI region J539.  
OS Mus musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=81054757; PubMed=6776525;  
RA Rudikoff S., Rao D.N., Gaudemans C.P.J., Potter M.;  
RT "Kappa Chain joining segments and structural diversity of antibody  
combining sites.";  
RT Proc. Natl. Acad. Sci. U.S.A. 77:4270-4274(1980).  
RN [2]  
RX X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
RX MEDLINE=88217852; PubMed=3449853;  
RA Suh S.W., Bhat T.N., Navia M.A., Cohen G.H., Rao D.N., Rudikoff S.,  
RA Davies D.R.;  
RT "The galactan-binding immunoglobulin Fab J539: an X-ray diffraction  
study at 2.6-A resolution.";  
RT Proteins 1:74-80(1986).  
CC -1 MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
BIND GALACTAN.  
CC PIR; A01942; KVSUS5.  
DR PDB; 2FBJ; 15-OCT-90.  
DR InterPro: IPR003006; Ig\_MHC.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; IgV; 1.  
KM Immunoglobulin V region; 3D-structure.  
FT DOMAIN 1 23 FRAMEWORK-1.  
FT DOMAIN 2 33 FRAMEWORK-2.  
FT DOMAIN 3 48 FRAMEWORK-3.  
FT DOMAIN 4 55 FRAMEWORK-4.  
FT DOMAIN 5 87 FRAMEWORK-5.  
FT DOMAIN 6 96 FRAMEWORK-6.  
FT DOMAIN 7 106 FRAMEWORK-7.  
FT DISULFID 23 87  
FT STRAND 4 7  
FT STRAND 10 14  
FT TURN 15 16  
FT STRAND 19 25  
FT STRAND 31 37  
FT TURN 39 40  
FT STRAND 44 48  
FT TURN 51 51  
FT STRAND 52 53  
FT STRAND 55 56  
FT TURN 59 60  
FT STRAND 61 66  
FT TURN 67 68  
FT STRAND 69 74  
FT HELIX 79 81  
FT STRAND 83 91  
FT TURN 92 93  
FT STRAND 94 97  
FT STRAND 101 106  
FT NON\_TER 107 107  
SQ SEQUENCE 107 AA; 11502 MW; EA30C9A3E903979C CRC64;

Query Match 71.2%; Score 396; DB 1; Length 107;  
Best Local Similarity 71.7%; Pred. No. 6, 4e-31;  
Matches 76; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

OY 1 QIVSOSPALISASFGKVMTCRASSSVNMHYOOKPCSSPKPMISATSNLSGVPAR 60  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 1 EIVLTOSPALISASFGKVTITCSASSVSLHWYOOKSGTSPKPMIYEISKLSGVPAR 60  
OY 61 FSGSGSTSYSLTISRVEDATYVCOQSSNPPFGGCTMLEIR 106  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 FSGSGSTSYSLTISRVEDATYVCOQNTYPLITFGAGTKLEIK 106

RESULT 13  
KV3J\_MOUSE  
ID KV3J\_MOUSE STANDARD: PRT: 111 AA.  
AC P01652;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 19 kappa chain V-III region ABPC 22/PC 9245.  
OS Mus musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE (ABPC 22).  
RX MEDLINE=79012520; PubMed=99744;  
RA McKean D.J., Bell M., Potter M.;  
RT "Mechanisms of antibody diversity: multiple genes encode structurally  
related mouse kappa variable regions.";  
RT Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).  
RN [2]  
RX MEDLINE=9073152; PubMed=103003;  
RA Weigert M., Garmalan L., Loh E., Schilling J., Hood L.E.;  
RT "Rearrangement of genetic information may produce immunoglobulin  
diversity.";  
RT Nature 276:785-790(1978).  
CC -1 MISCELLANEOUS: THE ABPC22 AND PC9241 SEQUENCES ARE IDENTICAL.  
CC PIR; A01935; KYSM6.  
DR HSSP: A01679; 2FBJ.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; IgV; 1.  
KM Immunoglobulin V region.  
FT DOMAIN 1 23 FRAMEWORK-1.  
FT DOMAIN 2 38 FRAMEWORK-2.  
FT DOMAIN 3 53 FRAMEWORK-3.  
FT DOMAIN 4 60 FRAMEWORK-4.  
FT DOMAIN 5 92 FRAMEWORK-5.  
FT DOMAIN 6 101 FRAMEWORK-6.  
FT DOMAIN 7 102 FRAMEWORK-7.  
FT DISULFID 23 92  
FT NON\_TER 111 111  
SQ SEQUENCE 111 AA; 12041 MW; D7DF0609303453CE CRC64;

Query Match 62.9%; Score 349.5; DB 1; Length 111;  
Best Local Similarity 63.6%; Pred. No. 1, 6e-26;  
Matches 70; Conservative 14; Mismatches 21; Indels 5; Gaps 1;

OY 2 IIVSOSPALISASFGKVTITCRASSV-----NYMHYOOKPCSSPKPMISATSNLSG 56  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 2 IIVLTOSPALISASFGKVTITCRASSVSDSYGNSFMHYOOKPCQPKLITLYLSNLSG 61

OY 57 VPARFSGSGTSTSLTISRVEDATYVCOQSSNPPFGGCTMLEIR 106  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 62 VPARFSGSGRTDFTLTIDPEADDAATYVCOQNNEDPYFGGCTKLEIK 111

RESULT 14  
KV3T\_MOUSE  
ID KV3T\_MOUSE STANDARD: PRT: 111 AA.  
AC P01672;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 19 kappa chain V-III region PC 7940.



Query Match	62.9%	Score 349.5;	DB 1;	Length 111;
Best Local Similarity	63.6%	Pred. No. 1.6e-26;		
Matches	70;	Conservative	10;	Mismatches 25;
			Indels	5;
			Gaps	1;

```

57 VPARESGSGSTSYSLTISRVEADATYYCQWSSNPPTFGGGTMEIR 106
||||| : : | | | | | ||||| : :
62 VPARESGSGSTDLNIHPVEEDAVTYQCQSRELPTFGGKTLEIK 111

```

T	1	23	FRAMEWORK-1.
T	24	38	COMPLEMENTARITY-DETERMINING-1.
T	39	53	FRAMEWORK-2.
T	54	60	COMPLEMENTARITY-DETERMINING-2.

Query Match	62.7%;	Score 348.5;	DB 1;	Length 111;
Best Local Similarity	62.7%;	Pred. No. 2e-26;		
Matches 69;	Conservative 15;	Mismatches 21;	Indels 5;	Gaps 1;

```

57 VPARFSGSGSTSYSLTISRVEADATYYCQOWSSNPPTFGGGMLEIR 106
    : ||||| |::| | ||||| |::| ||||| |::|
62 IPARFSGSGSDFTLNHPVEEDATYYCQGSNEDPYTFGGGIKLEIK 111

```

Search completed: November 27, 2002, 07:18:57  
Job time : 16.1429 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 06:27:07 ; Search time 58.9344 Seconds  
(without alignments)  
370.599 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 556

Sequence: 1 QIVLSQSPALISAPGKVT.....CQKMSNPFGCGTMLEIR 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	455.5	81.9	97	11 Q9JL76	Q9JL76 mus musculu
2	453	81.5	134	11 Q8VDD0	Q8VDD0 mus musculu
3	441	79.3	235	11 Q91W12	Q91W12 mus musculu
4	427	76.8	106	5 Q9U410	Q9U410 schistosoma
5	368.5	66.3	101	11 Q9JL78	Q9JL78 mus musculu
6	352	63.3	109	4 Q9U178	Q9U178 homo sapien
7	343.5	61.8	111	11 Q920E9	Q920E9 mus musculu
8	341.5	61.4	108	4 Q9UL70	Q9UL70 homo sapien
9	339.5	61.1	108	4 Q9UL79	Q9UL79 homo sapien
10	331	59.5	107	4 Q96SA9	Q96SA9 homo sapien
11	327	58.8	107	4 Q9UL81	Q9UL81 homo sapien
12	323.5	58.2	108	4 Q9UL77	Q9UL77 homo sapien
13	323	58.1	109	4 Q9UL85	Q9UL85 homo sapien
14	319	57.4	109	4 Q9UL86	Q9UL86 homo sapien
15	317.5	57.1	109	11 Q920E6	Q920E6 mus musculu
16	317.5	57.1	234	11 Q8R062	Q8R062 mus musculu

17	317.5	57.1	298	11 Q90YF0	Q90YF0 mus musculu
18	315.5	56.7	108	4 Q9UL83	Q9UL83 homo sapien
19	313.5	56.4	233	11 Q91WS9	Q91WS9 mus musculu
20	309.5	55.7	214	11 Q91RL5	Q91RL5 mus musculu
21	309	55.6	238	11 Q8VC16	Q8VC16 mus musculu
22	307.5	55.3	127	11 Q925S9	Q925S9 mus musculu
23	307.5	55.3	234	11 Q91W18	Q91W18 mus musculu
24	306.5	55.1	108	11 Q8VJ10	Q8VJ10 mus musculu
25	299.5	53.9	107	11 Q9JL84	Q9JL84 mus musculu
26	296.5	53.3	103	11 Q9JL80	Q9JL80 mus musculu
27	293.5	52.5	234	11 Q8VCP0	Q8VCP0 mus musculu
28	292	52.5	238	11 Q99M37	Q99M37 mus musculu
29	285.5	51.3	99	11 Q9JL74	Q9JL74 mus musculu
30	281.5	50.6	107	11 Q9ER29	Q9ER29 mus musculu
31	280.5	50.4	116	4 Q96PF6	Q96PF6 homo sapien
32	280	50.4	104	11 Q9JL82	Q9JL82 mus musculu
33	280	50.4	239	11 Q8VCS5	Q8VCS5 mus musculu
34	280	50.4	241	11 Q921A6	Q921A6 mus musculu
35	277	49.8	239	4 Q8TCD0	Q8TCD0 homo sapien
36	270.5	48.7	114	4 Q9UL80	Q9UL80 homo sapien
37	265.5	47.8	109	6 Q9N0W5	Q9N0W5 oryctolagus
38	262	47.5	237	4 Q8WUK4	Q8WUK4 homo sapien
39	262	47.1	237	4 Q8WTU6	Q8WTU6 homo sapien
40	261.5	47.0	234	11 Q8R028	Q8R028 mus musculu
41	242.5	43.6	236	4 Q96E61	Q96E61 homo sapien
42	235.5	42.4	129	11 Q8VDE2	Q8VDE2 mus musculu
43	235	42.3	108	4 Q96S80	Q96S80 homo sapien
44	235	42.3	110	4 Q8TE63	Q8TE63 homo sapien
45	234.5	42.2	112	4 Q96JDI	Q96JDI homo sapien

## ALIGNMENTS

RESULT 1				
ID	Q9JL76	PRELIMINARY:	PR:	97 AA.
AC	Q9JL76:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Anti-myosin immunoglobulin light chain variable region (Fragment).			
DE	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN-DBA/2:			
RC	MEDLINE-20448942; PubMed-10992488;			
RA	Maikiel S., Liao L., Cunningham M.W., Diamond B.;			
RT	"T-Cell-dependent antibody response to the dominant epitope of streptococcal polysaccharide, N-acetyl-galactosamine, is cross-reactive with cardiac myosin."			
RT	Infect. Immun. 68:5803-5808(2000).			
RU	EMBL; AF206030; AAF69328.1; --			
DR	HSSP; P01679; 2EBJ.			
DR	InterPro; IPR003006; IG_MHC.			
DR	InterPro; IPR003596; IG_V.			
DR	Pfam; PF00047; 1g; 1.			
DR	SMART; SM00406; IGV; 1.			
FT	NON-TER			
FT	NON-TER			
SO	SEQUENCE			
Query Match				
Best Local Similarity 81.9%; Score 455.5; DB 11; Length 97;				
Matches 89; Conservative 91.8%; Pred. NO. 1.1e-38;				
Mismatches 2; Indels 5; Gaps 1;				
OY 11 LSASGKVTMTCRASSVYVHWYQQRGSSPKPMISNTSLASGVAPRFSGSGGSSTSY 70				
DB 1 LSASGKVTMTCRASSVYVHWYQQRGSSPKPMIYATSLASGVAPRFSGSGGSSTSY 60				

```

QY 71 SLTISRVEADATYYCCQWSSN-PPYFGGTMLEIR 106
DB 61 SLTISRVEADATYYCCQWSSN-PPYFGGTMLEIR 97

RESULT 2
OBVDDO
ID 08VDDO PRELIMINARY: PRT: 134 AA.
AC 08VDDO:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Anti-MOG 212 variable light chain (Fragment).
GN ANTI-MOG KAPPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Cheralovsky Y.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Semi P.;
RT "Targeting T cells to the CNS.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AA416331; CAC94866.1; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00409; Ig_1.
DR SMART: SM00406; IGV; 1.
DR NON_TER 134
SQ SEQUENCE 134 AA; 14525 MW; CFDF8E2236E2D0CF CRC64;

Query Match 81.5%; Score 453; DB 11; Length 134;
Best Local Similarity 83.0%; Pred. No. 3e-38;
Matches 88; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 QIVLSOSPALISAPGEKVTMTCRASSVYVMHMYOQKPGSSPKPWISATSNLASGVPAR 60
DB 23 QIVLTSPALMSASPGKVTMTCSASSISYMHMYOQKPGTSPKRWLYDTSKLASGVPAR 82

QY 61 FSSGSGTSTSLTISRVEADATYYCCQWSSN-PPYFGGTMLEIR 106
DB 83 FSSGSGTSTSLTISRVEADATYYCHORSYPWTFGGTKLEIK 128

RESULT 3
Q91W12 PRELIMINARY: PRT: 235 AA.
AC Q91W12:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:65382).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST TUMOR;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC006643; AAH06643.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_2.

```

```

DR PROSITE: PS00290; IG_MHC; UNKNOWN.1.
SQ SEQUENCE 235 AA; 26021 MW; 5FC73BDEBD5E8FEF CRC64;

Query Match 79.3%; Score 441; DB 11; Length 235;
Best Local Similarity 78.3%; Pred. No. 9.5e-37;
Matches 83; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 QIVLSOSPALISAPGEKVTMTCRASSVYVMHMYOQKPGSSPKPWISATSNLASGVPAR 60
DB 23 QIVLTSPALMSASPGKVTMTCSASSISYMHMYOQKPGTSPKRWLYDTSKLASGVPAR 82

QY 61 FSSGSGTSTSLTISRVEADATYYCCQWSSN-PPYFGGTMLEIR 106
DB 83 FSSGSGTSTSLTISRVEADATYYCCQWSSN-PPYFGGTMLEIR 128

RESULT 4
Q9U410 PRELIMINARY: PRT: 106 AA.
AC Q9U410:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Monoclonal anti-Idiotypic antibody NP30 immunoglobulin light chain
variable region (Fragment).
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoidae;
Schistosomatidae; Schistosoma.
NCBI_TaxID=6182;
RN (1)
RP SEQUENCE FROM N.A.
RA Song X.T., Feng Z.Q., Qiu Z.N., Li Y.O., Huang H.L., Guan X.H.;
RT "Amplification, cloning and sequence analysis of the light chain
variable region gene of monoclonal anti-Idiotypic antibody NP30 of
Schistosoma japonicum.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF207620; AAJ19434.1; -.
DR HSSP: P01679; 2FBI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 106
SQ SEQUENCE 106 AA; 11478 MW; F20F544426BAE63E CRC64;

Query Match 76.8%; Score 427; DB 5; Length 106;
Best Local Similarity 78.8%; Pred. No. 9.5e-36;
Matches 82; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 3 VLSQSPALISAPGEKVTMTCRASSVYVMHMYOQKPGSSPKPWISATSNLASGVPAR 62
DB 3 LITQSPALMSASPGKVTMTCSASSISYVMHMYOQKPGTSPKRWLYDTSKLASGVPAR 82

QY 63 GSSGSGTSTSLTISRVEADATYYCCQWSSN-PPYFGGTMLEIR 106
DB 63 GSSGSGTSTSLTISRVEADATYYCCQWSSN-PPYFGGTMLEIR 106

RESULT 5
Q9UL78 PRELIMINARY: PRT: 101 AA.
AC Q9UL78:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Anti-myosin immunoglobulin light chain variable region
(fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)

```



[illegible]

```
RT      "Molecular analysis of polyreactive monoclonal antibodies from  
KT      rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin  
RT      antibody V region genes.";  
RL      J. Immunol. 161:2020-2031(1998).  
DR      EMBL; U96396; AAB68785.1; -.  
DR      InterPro: IPR003006; Ig_MHC.  
FT      Pfam: PF00047; 1g; 1.  
FT      NON_TER  
SO      SEQUENCE          107 AA; 11520 MW; 4BB43E9C5B577F16 CMC6d;  
  
Query Match           59.5%; Score 311; DB 4; Length 107;  
Best Local Similarity 62.3%; Pred. No. 4.7e-26;  
Matches    66; Conservative   20; Mismatches   18; Indels     2; Gaps       2.
```

```
OY      2 IVLSGSPALILASPEEKYTMTCRASSV-NYNHMYQOKRGSSPKPMISATSNLASGVPAR 60  
|::|||:||||:|||:|||||:-|-:|||||:-|||:||||:|  
Db      2 IQMGTPSPLSLASVDGYITTCRASQSISTSYLNWYQKGRKPMLIIYAASSLGSGVP SR 61  
| | || || : :: || : - | | | | | | | | : | | | | : |:  
OY      61 FSGSGSTYSLTISRVEADDAITYYCOWMSNPTFGGGTMLEIR 106  
| | | | | | : | | | | | | | | : | | | | | : |:  
Db      62 FSGSGSGTDFTLTISGLQAEDPATYYCOO-SYSTLTFFGGTGKEIK 106  
| | | | | | : | | | | | | | | : | | | | | : |:  
  
RESULT 11  
O9UL81 PRELIMINARY; PRT; 107 AA.  
ID O9UL81;  
AC O9UL81;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DI 01-MAY-2000 (TrEMBRel. 13, Last sequence update)  
DE 01-DEC-2001 (TrEMBRel. 19, Last annotation update)  
DEF Myosin-reactive immunoglobulin light chain variable region  
DE (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
CX NCBI_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98277139; PubMed=9614934;  
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,  
RA Young D.C.;  
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
RT fetus."  
KL Clin. Immunol. Immunopathol. 87:184-192(1998).  
RL DR EMBL; AF035033; AADB6269.1; -.  
DR HSSP; POLI607; IREI.  
DR InterPro: IPR003006; Ig_MHC.  
DR InterPro: IPR003596; Ig_V.  
DR Pfam; PF00047; 1g; 1.  
DR SMART; SM00406; IGv; 1.  
FT NON_TER  
FT NON_TER 1  
FT NON_TER 107  
SQ SEQUENCE 107 AA; 11501 MW; 070549FE5D0754748 CMC6d;
```

DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE Myosin-reactive immunoglobulin light chain variable region  
DE (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98277139; PubMed=9614934;  
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,  
RA Young D.C.;  
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
fetus".  
RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
DR EMBL; AF035037; AAD56273.1; -.  
DR HSP; P01607; IREL.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART; SM00406; IgV; 1.  
FT NON\_TER 1  
FT SEQUENCE 108 AA; 11738 MW; C06681716CADI6F3 CRC64;  
SQ  
Query Match 58.2%; Score 323.5; DB 4; Length 108;  
Best Local Similarity 60.4%; Pred. No. 2.7e-25;  
Matches 64; Conservative 19; Mismatches 22; Indels 1; Gaps 1;  
QY 2 IVLSOSPALISAPGEKVTMTCRASSV--NYMHVYQKRGSSPKWISATSNLASGVP 60  
DB 2 IGMTSPSLASVGDRAVITTCRASQISISLYMWYQKRGKAPNLLIYASSLSQSVSR 61  
QY 61 FSGSGSGTSYSLTISRVEADATYTCQOMSSNPPTFGGTMLEIR 106  
DB 62 FSGSGSGTDFLTLSLQPEDFATYTCQOSYSTSMFTGCTKVEIK 107  
RESULT 13  
Q9UL85 PRELIMINARY; PRT; 109 AA.  
AC Q9UL85;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE Myosin-reactive immunoglobulin kappa chain variable region  
DE (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98277139; PubMed=9614934;  
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,  
RA Young D.C.;  
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
fetus".  
RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
DR EMBL; AF035029; AAD56265.1; -.  
DR HSP; P80362; IRTL.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IgV; 1.  
FT NON\_TER 1  
FT SEQUENCE 109 AA; 11761 MW; FBIE43E7C7AFACCC CRC64;  
SQ  
Query Match 58.1%; Score 323; DB 4; Length 109;  
Best Local Similarity 57.4%; Pred. No. 3.1e-25;  
Matches 62; Conservative 23; Mismatches 21; Indels 2; Gaps 2;

QY 1 QIVLSOSPALISAPGEKVTMTCRASSV--NYMHVYQKRGSSPKWISATSNLASGVP 59  
DB 1 EIVTOSPATLSVSPERATILSCMASQISSINLAMYQKPGQAPRLIYGASTRATGIPA 60  
QY 60 FSGSGSGTSYSLTISRVEADATYTCQOMSSNPPTFGGTMLEIR 106  
DB 61 FSGSGSGTEFTLTLSLQSEDPATYTCQOSYSTSMFTGCTKVEIK 108  
RESULT 14  
Q9UL86 PRELIMINARY; PRT; 109 AA.  
AC Q9UL86;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE Myosin-reactive immunoglobulin kappa chain variable region  
DE (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98277139; PubMed=9614934;  
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,  
RA Young D.C.;  
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
fetus".  
RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
DR EMBL; AF035028; AAD56264.1; -.  
DR HSP; P80362; IRTL.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IgV; 1.  
FT NON\_TER 1  
FT SEQUENCE 109 AA; 11928 MW; 243325F2C7DAC83 CRC64;  
SQ  
Query Match 57.4%; Score 319; DB 4; Length 109;  
Best Local Similarity 57.4%; Pred. No. 7.8e-25;  
Matches 62; Conservative 20; Mismatches 24; Indels 2; Gaps 1;  
QY 1 QIVLSOSPALISAPGEKVTMTCRASSV--NYMHVYQKRGSSPKWISATSNLASGVP 58  
DB 1 EIVTOSPATLSVSPERATILSCMASQISSINLAMYQKPGQAPRLIYGASTRATGIP 60  
QY 59 ARFSGSGTSYSLTISRVEADATYTCQOMSSNPPTFGGTMLEIR 106  
DB 61 DRFSGSGTDFLTLSLQPEDFATYTCQOSYSTSMFTGCTKVEIK 108  
RESULT 15  
Q920E6 PRELIMINARY; PRT; 109 AA.  
AC Q920E6;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
DE pterin-mimicking anti-idiotypic kappa chain variable region  
DE (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Atkin J.D., Iape A., Jennings I.G., Horvath O., Cotton R.G.H.;  
RT "Definition of the idiotype of Pterin-Mimicking Antibodies Expressed  
in Mammalian Cells".  
Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF307938; AL09422.1; -  
DR InterPro; IPR03006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 1.  
FT NON\_TER 1 1  
SO SEQUENCE 109 AA; 11943 MW; DAD3F98E05DD1501 CRC64;

Query Match 57.1%; Score 317.5; DB 11; Length 109;  
Best Local Similarity 59.4%; Pred. No. 1.1e-24;  
Matches 63; Conservative 15; Mismatches 27; Indels 1; Gaps 1;

OY 2 IYLSOSPAILSPGEXVTMTCRASSV-NYMHYQOKFGSSPKFWISATSNLASGVPAR 60  
|::||||| ||| |||::||||| ::||: ||||| | ||: | |||||  
Db 2 IQMTQSPASLSASVGEVTITCRASGNHNYLAWYQOKGKSPQLLYNNAKTLADGVPSR 61  
OY 61 FSGSGSGTYSILTISRVEAEDATYYCOQSSNPPTFGGTMLEIR 106  
||||||| |||::||: ||||| : ||| ||||| |||  
Db 62 FSGSGSGTQYSLKINSLOPEDFGSYCOHFWSTFWTFGGTKLEIK 107

Search completed: November 27, 2002, 07:22:55  
Job time : 60.9344 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:17:51 : Search time 19.2355 Seconds

(without alignments)  
162.139 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 556

Sequence: 1 QIVLSQSPALISASPGKVT.....CQKWSNPRTGCGTMEIR 106

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfilest1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	526	94.6	106	3	US-08-783-853A-105
2	526	94.6	106	4	US-09-344-050-105
3	526	94.6	107	3	US-08-783-853A-11
4	526	94.6	107	4	US-09-344-050-11
5	526	94.6	112	3	US-08-783-853A-103
6	526	94.6	112	4	US-09-344-050-103
7	521	93.7	128	1	US-08-476-275-4
8	521	93.7	128	4	US-08-476-275-4
9	507	91.2	129	2	US-08-449-287-2
10	507	91.2	129	4	US-09-423-439-18
11	507	91.2	235	4	US-09-423-439-58
12	507	91.2	235	4	US-09-011-769A-23
13	490	88.1	105	1	US-08-459-310-4
14	486	87.4	107	1	US-08-211-202-3
15	486	87.4	246	1	US-08-469-486-57
16	486	87.4	246	2	US-08-469-658-57
17	484	87.1	252	3	US-08-279-772A-6
18	484	87.1	252	4	US-08-902-486-9
19	482	86.7	237	2	US-08-468-252-5
20	482	86.7	237	3	US-08-668-706B-5
21	482	86.7	237	5	PCT-US95-10740-5
22	475	85.4	97	2	US-08-308-494A-11
23	473.5	85.2	110	3	US-08-836-561-33
24	471.5	84.8	105	1	US-08-211-202-110
25	469	84.4	97	4	US-08-280-028-4
26	466	83.8	129	4	US-08-116-778E-2
27	466	83.8	129	2	US-08-438-562-2

28	466	83.8	129	2	US-08-483-528B-92	Sequence 92, Appl
29	464	83.5	128	4	US-08-619-491-2	Sequence 2, Appl1
30	464	83.5	128	5	PCT-US95-07302-2	Sequence 2, Appl1
31	462	83.1	128	2	US-08-656-586-2	Sequence 2, Appl1
32	457	82.2	222	2	US-08-190-199A-67	Sequence 67, Appl
33	457	82.2	235	2	US-08-190-199A-61	Sequence 61, Appl
34	456	82.0	281	4	US-09-423-439-44	Sequence 44, Appl
35	456	82.0	666	4	US-09-423-439-51	Sequence 51, Appl
36	455	81.8	106	3	US-08-397-411-9	Sequence 9, Appl1
37	455	81.8	129	3	US-08-783-853A-99	Sequence 99, Appl
38	455	81.8	129	4	US-09-344-050-99	Sequence 99, Appl
39	454	81.7	125	3	US-08-783-853A-78	Sequence 78, Appl
40	454	81.7	125	4	US-09-344-050-78	Sequence 78, Appl
41	453	81.5	127	1	US-08-483-882-80	Sequence 80, Appl
42	453	81.5	127	2	US-08-483-389-80	Sequence 80, Appl
43	453	81.5	127	2	US-08-487-113D-80	Sequence 80, Appl
44	453	81.5	127	2	US-08-473-503-80	Sequence 80, Appl
45	453	81.5	127	2	US-08-483-932-80	Sequence 80, Appl

## ALIGNMENTS

RESULT 1  
US-08-783-853A-105  
; Sequence 105 Application US/08783853A  
; Patent No. 6005091  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; APPLICANT: Church, William  
; APPLICANT: Gross, Mitchell  
; APPLICANT: Feuerstein, Gloria  
; APPLICANT: Nichols, Andrew  
; APPLICANT: Padlan, Eduardo  
; APPLICANT: Patel, Arunbhai  
; APPLICANT: Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/783, 853A  
; FILING DATE: 16-JAN-1997  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 60/029,119  
; FILING DATE: 24-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX:  
; TELETYPE:  
; INFORMATION FOR SEQ ID NO: 105:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
US-08-783-853A-105

Query Match  
Best Local Similarity 94.6%; Score 526; DB 3; Length 106;  
Pred. No. 1.6e-42;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 QIVLSOSPALLSAPGKVTMTTCRASSSVNYMHYQOKPSSSPKPVISATSNLASGVPAR 60  
DB 1 QIVLSOSPALLSAPGKVTMTTCRASSSVNYMHYQOKPSSSPKPVISATSNLASGVPAR 60

QY 61 FSSGSGTSTSLTISRVEDADATYYCOQWSSNPRTGGGTMLEIR 106  
DB 61 FSSGSGTSTSLTISRVEDADATYYCOQWSSNPRTGGGTMLEIR 106

## RESULT 2

US-09-344-050-105  
Sequence 105, Application US/09344050  
Patent No. 6391299

## GENERAL INFORMATION:

APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Giora  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,050  
FILING DATE: 24-JUN-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:

TELEX:  
INFORMATION FOR SEQ. ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:

US-09-344-050-105

Query Match  
Best Local Similarity 94.6%; Score 526; DB 4; Length 106;  
Pred. No. 1.6e-42;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 QIVLSOSPALLSAPGKVTMTTCRASSSVNYMHYQOKPSSSPKPVISATSNLASGVPAR 60  
DB 1 QIVLSOSPALLSAPGKVTMTTCRASSSVNYMHYQOKPSSSPKPVISATSNLASGVPAR 60

QY 61 FSSGSGTSTSLTISRVEDADATYYCOQWSSNPRTGGGTMLEIR 106  
DB 61 FSSGSGTSTSLTISRVEDADATYYCOQWSSNPRTGGGTMLEIR 106

## RESULT 3

US-08-783-853A-11  
Sequence 11, Application US/08783853A  
Patent No. 6005091

## GENERAL INFORMATION:

APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Giora  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/783,853A  
FILING DATE: 16-JAN-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/029,119  
FILING DATE: 24-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:

TELEX:  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
US-08-783-853A-11

Query Match  
Best Local Similarity 94.6%; Score 526; DB 3; Length 107;  
Pred. No. 1.7e-42;

Matches 101: Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QIVLSQSPAILSASPGKVTMTCRASSSVNMYHWYQOKRGSSPKPWISATSNLASGVPAR 60  
|||||  
Db 1 QIVLSQSPAILSASPGKVTMTCRASSSVNMYHWYQOKRGSSPKPWISATSNLASGVPAR 60

Qy 61 FSGSGSTSYSLTISRVEADATYCCOOWSNPTFGGCTMLEIR 106  
|||||  
Db 61 FSGSGSTSYSLTISRVEADATYCCOOWSNPTFGGCTMLEIR 106

RESULT 4  
US-09-344-050-11

; Sequence 11, Application US/09344050

; Patent No. 6391299

; GENERAL INFORMATION:

; APPLICANT: Blackburn, Michael

; APPLICANT: Church, William

; APPLICANT: Gross, Mitchell

; APPLICANT: Feuerstein, Gloria

; APPLICANT: Nichols, Andrew

; APPLICANT: Padlan, Eduardo

; APPLICANT: Patel, Arundhal

; APPLICANT: Sylvester, Daniel

; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/344,050

; FILING DATE: 24-JUN-1999

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/783,853

; FILING DATE: 16-JAN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Baumelster, Kirk

; REGISTRATION NUMBER: 33,833

; REFERENCE/DOCKET NUMBER: P50438

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-5096

; TELEFAX:

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 107 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: Internal

; ORIGINAL SOURCE:

; US-09-344-050-11

Query Match 94.6%; Score 526; DB 4; Length 107;

Best Local Similarity 95.3%; Pred. No. 1.7e-42;

Matches 101: Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QIVLSQSPAILSASPGKVTMTCRASSSVNMYHWYQOKRGSSPKPWISATSNLASGVPAR 60  
|||||  
Db 1 QIVLSQSPAILSASPGKVTMTCRASSSVNMYHWYQOKRGSSPKPWISATSNLASGVPAR 60

Db 1 QIVLSQSPAILSASPGKVTMTCRASSSVNMYHWYQOKRGSSPKPWISATSNLASGVPAR 60

Qy 61 FSGSGSTSYSLTISRVEADATYCCOOWSNPTFGGCTMLEIR 106  
|||||  
Db 61 FSGSGSTSYSLTISRVEADATYCCOOWSNPTFGGCTMLEIR 106

RESULT 5  
US-08-783-853A-103

; Sequence 103, Application US/08783853A

; Patent No. 6005091

; GENERAL INFORMATION:

; APPLICANT: Blackburn, Michael

; APPLICANT: Church, William

; APPLICANT: Gross, Mitchell

; APPLICANT: Feuerstein, Gloria

; APPLICANT: Nichols, Andrew

; APPLICANT: Padlan, Eduardo

; APPLICANT: Patel, Arundhal

; APPLICANT: Sylvester, Daniel

; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/783,853A

; FILING DATE: 16-JAN-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/029,119

; FILING DATE: 24-OCT-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Baumelster, Kirk

; REGISTRATION NUMBER: 33,833

; REFERENCE/DOCKET NUMBER: P50438

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-5096

; TELEFAX:

; INFORMATION FOR SEQ ID NO: 103:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 112 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: Internal

; ORIGINAL SOURCE:

; US-08-783-853A-103

Query Match 94.6%; Score 526; DB 3; Length 112;

Best Local Similarity 95.3%; Pred. No. 1.7e-42;

Matches 101: Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QIVLSQSPAILSASPGKVTMTCRASSSVNMYHWYQOKRGSSPKPWISATSNLASGVPAR 60  
|||||  
Db 1 QIVLSQSPAILSASPGKVTMTCRASSSVNMYHWYQOKRGSSPKPWISATSNLASGVPAR 60

Db 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 106

## RESULT 6

US-09-344-050-103

; Sequence 103, Application US/09344050  
; Patent No. 6391299

## GENERAL INFORMATION:

; APPLICANT: Blackburn, Michael  
; APPLICANT: Church, William  
; APPLICANT: Gross, Mitchell  
; APPLICANT: Feuerstein, Gloria  
; APPLICANT: Nichols, Andrew  
; APPLICANT: Padlan, Eduardo  
; APPLICANT: Patel, Arunbhai  
; APPLICANT: Sylvester, Daniel

; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; TITLE OF INVENTION: OF THROMBOSIS

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/344,050  
; FILING DATE: 24-JUN-1999

## CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/783,853

; FILING DATE: 16-JAN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Baumeister, Kirk

; REGISTRATION NUMBER: 33,833

; REFERENCE/DOCKET NUMBER: P50438

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-5096

; TELEFAX:

; INFORMATION FOR SEQ ID NO: 103:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 112 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: Internal

; ORIGINAL SOURCE:

; US-09-344-050-103

Query Match 94.6%; Score 526; DB 4; Length 112;  
Best Local Similarity 95.3%; Pred. No. 1.7e-42;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QIVLSQSPALLSAPGKVTMTCTCRASSSVNYMHYQOKPCSSPKPWISATSNLASGVPAR 60  
Db 1 QIVLSQSPALLSAPGKVTMTCTCRASSSVNYMHYQOKPCSSPKPWISATSNLASGVPAR 60

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 106

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 106

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 106

## RESULT 7

US-08-476-275-4  
; Sequence 4, Application US/08476275  
; Patent No. 5776456

## GENERAL INFORMATION:

; APPLICANT: Anderson, Darrell R.  
; APPLICANT: Hanna, Nabil  
; APPLICANT: Leonard, John E.  
; APPLICANT: Newman, Roland A.  
; APPLICANT: Reff, Mitchell E.  
; APPLICANT: Rastetter, William H.

; TITLE OF INVENTION: Therapeutic Application of Chimeric and  
; TITLE OF INVENTION: Radiolabeled Antibodies to Human B Lymphocyte Restricted  
; TITLE OF INVENTION: Differentiation Antigen for the Treatment of B-Cell  
; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
; STREET: 699 Prince St.  
; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22314

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/476,275  
; FILING DATE: 07-JUN-1995

## CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/149,099  
; FILING DATE: 03-NOV-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/978,891  
; FILING DATE: 13-NOV-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Teskin, Robin L.

; REGISTRATION NUMBER: 35,030

; REFERENCE/DOCKET NUMBER: 012712-155

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-836-6620

; TELEFAX: 703-836-2021

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 128 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-476-275-4

Query Match 93.7%; Score 521; DB 1; Length 128;  
Best Local Similarity 92.5%; Pred. No. 5.9e-42;  
Matches 98; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QIVLSQSPALLSAPGKVTMTCTCRASSSVNYMHYQOKPCSSPKPWISATSNLASGVPAR 60  
Db 23 QIVLSQSPALLSAPGKVTMTCTCRASSSVNYMHYQOKPCSSPKPWISATSNLASGVPAR 60

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

Qy 61 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIR 106  
Db 83 FSSGSGTSLTISRVEDAATYYCQOWSINPRTFGGKTLEIK 128

APPLICANT: NEMMAN, ROLAND A.  
APPLICANT: REEF, MITCHELL E.  
APPLICANT: RASTETTER, WILLIAM H.  
TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND  
TITLE OF INVENTION: RADIOLABELLED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED  
TITLE OF INVENTION: DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL  
TITLE OF INVENTION: LYMPHOMA  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY WINTHROP  
STREET: 1100 New York Avenue, N.W., Ninth FL.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,815B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/978,891  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 23522-0157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ. ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 128 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-475-815B-7

Query Match 93.7%; Score 521; DB 4; Length 128;  
Best Local Similarity 92.5%; Pred. No. 5,9e-42;  
Matches 98; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIVLSQSPALISASPEKYMTCRASSSVNYMHYQOKPSSPKPMISATSNLASGVPR 60  
DB 23 QIVLSQSPALISASPEKYMTCRASSSVYIHWFQOKPSSPKPMIYATSNLASGVPR 82

QY 61 FSGSGSGTSLTISRVEADATYYCOQWSSNPFTFGGTMLEIR 106  
DB 83 FSGSGSGTSLTISRVEADATYYCOQWTSNPFTFGGTMLEIR 128

RESULT 9  
US-08-449-287-2  
Sequence 2, Application US/08449287  
Patent No. 5877293  
GENERAL INFORMATION:  
APPLICANT: ADAIR, John Robert  
APPLICANT: BODMER, Mark William  
APPLICANT: MOUNTAIN, Andrew  
APPLICANT: OWENS, Raymond John  
TITLE OF INVENTION: CDR Grafted Anti-CEA Antibodies and  
TITLE OF INVENTION: Their Production  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500

CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,287  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/154,389  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT GB91/01108  
FILING DATE: 05-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9014932.9  
FILING DATE: 05-JUL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT GB90/02017  
FILING DATE: 21-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 40283/110 CARA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 129 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-449-287-2

Query Match 91.2%; Score 507; DB 2; Length 129;  
Best Local Similarity 91.5%; Pred. No. 1.2e-40;  
Matches 97; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIVLSQSPALISASPEKYMTCRASSSVNYMHYQOKPSSPKPMISATSNLASGVPR 60  
DB 23 QIVLSQSPALISASPEKYMTCRASSSVYIHWFQOKPSSPKPMIYATSNLASGVPR 82

QY 61 FSGSGSGTSLTISRVEADATYYCOQWSSNPFTFGGTMLEIR 106  
DB 83 FSGSGSGTSLTISRVEADATYYCOHWSKPTFGGTMLEIR 128

RESULT 10  
US-09-423-439-18  
Sequence 18, Application US/09423439  
Patent No. 6339070  
GENERAL INFORMATION:  
APPLICANT: EMERY, Stephen Charles  
APPLICANT: BLAKEY, David Charles  
TITLE OF INVENTION: CHEMICAL COMPOUNDS  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Winthrop, L.L.P.  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/423.439  
FILING DATE: 09-NOV-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB98/01294  
FILING DATE: 05-MAY-1998  
APPLICATION NUMBER: GB 9709421.3  
FILING DATE: 10-MAY-1997  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 235 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-423-439-18

Query Match 91.2%; Score 507; DB 4; Length 235;  
Best Local Similarity 91.5%; Pred. No. 2,4e-40;  
Matches 97; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIVLSQSPALLSASGKGVMTTCRASSSVNTHMYQOKPGSSPKPWISATSNLASGVAR 60  
DB 23 QTVLSQSPALLSASGKGVMTTCRASSSVTYIHWYQOKPGSSPKSWIATSNLASGVAR 82  
QY 61 FSGSGSGTSLTISRVEAEDAATYYCOQWSSNPPTFGGTMLEIR 106  
DB 83 FSGSGSGTSLTISRVEAEDAATYYCOHWSKPPTEGGTKLEIK 128

RESULT 11  
US-09-423-439-58  
Sequence 58, Application US/09423439  
Patent No. 6339070  
GENERAL INFORMATION:  
APPLICANT: EMERY, Stephen Charles  
TITLE OF INVENTION: CHEMICAL COMPOUNDS  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Winthrop, L.L.P.  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/423.439  
FILING DATE: 09-NOV-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB98/01294  
FILING DATE: 05-MAY-1998  
APPLICATION NUMBER: GB 9709421.3  
FILING DATE: 10-MAY-1997  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 235 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 58:  
US-09-423-439-58

Query Match 91.2%; Score 507; DB 4; Length 235;

Best Local Similarity 91.5%; Pred. No. 2,4e-40;  
Matches 97; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIVLSQSPALLSASGKGVMTTCRASSSVNTHMYQOKPGSSPKPWISATSNLASGVAR 60  
DB 23 QTVLSQSPALLSASGKGVMTTCRASSSVTYIHWYQOKPGSSPKSWIATSNLASGVAR 82  
QY 61 FSGSGSGTSLTISRVEAEDAATYYCOQWSSNPPTFGGTMLEIR 106  
DB 83 FSGSGSGTSLTISRVEAEDAATYYCOHWSKPPTEGGTKLEIK 128

RESULT 12  
US-09-011-769A-23  
Sequence 23, Application US/09011769A  
Patent No. 6436691  
GENERAL INFORMATION:  
APPLICANT: SLATER, Anthony M.  
BLAKEY, David C.  
DAVIES, David H.  
HENNAM, John F.  
HENNEQUIN, Laurent F.A.  
MARSHAM, Peter R.  
DOWELL, Robert I.  
TITLE OF INVENTION: Chemical Compounds  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Madison & Sutro, LLP  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 Mb disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/011.769A  
FILING DATE: 13-Feb-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB96/01975  
FILING DATE: 13-AUG-1996  
APPLICATION NUMBER: GB 9612295.7  
FILING DATE: 12-JUN-1996  
APPLICATION NUMBER: GB 9611019.2  
FILING DATE: 25-MAY-1996  
APPLICATION NUMBER: GB 9516810.0  
FILING DATE: 16-AUG-1995  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 235 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-09-011-769A-23

Query Match 91.2%; Score 507; DB 4; Length 235;  
Best Local Similarity 91.5%; Pred. No. 2,4e-40;  
Matches 97; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIVLSQSPALLSASGKGVMTTCRASSSVNTHMYQOKPGSSPKPWISATSNLASGVAR 60  
DB 23 QTVLSQSPALLSASGKGVMTTCRASSSVTYIHWYQOKPGSSPKSWIATSNLASGVAR 82  
QY 61 FSGSGSGTSLTISRVEAEDAATYYCOQWSSNPPTFGGTMLEIR 106  
DB 83 FSGSGSGTSLTISRVEAEDAATYYCOHWSKPPTEGGTKLEIK 128

```

RESULT 13
US-08-459-310-4
; Sequence 4, Application US/08459310
; Patent No. 5645817
GENERAL INFORMATION:
APPLICANT: Seemann, Gerhard
APPLICANT: Bosslet, Klaus
TITLE OF INVENTION: Granulocyte-Binding Antibody Constructs,
TITLE OF INVENTION: Their Preparation and Use
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flunegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,310
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/100,963
FILING DATE: 03-AUG-1993
DE P 422 58 53.7
APPLICATION NUMBER: DE P 422 58 53.7
FILING DATE: 05-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wadler, Linda A.
REGISTRATION NUMBER: 33,218
REFERENCE/DOCKET NUMBER: 02481.1317-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-459-310-4

Query Match      88.1%; Score 490; DB 1; Length 105;
Best Local Similarity 91.3%; Pred. No. 3.8e-39;
Matches    95; Conservative   2; Mismatches     7; Indels       0; Gaps        0;

QY          2 IVLSGPAILLASPEKEXTMTCRASSVNYVMWYOQKFGSSPKPMIATSNLASGVPARF 61
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         2 IQLVSPAILLASEPEKXTMTCRASSVSYNHMWYOQKFESSPKPMIVATSNLASGVPARF 61
||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY          62 SGSGSGTSYSLLTIRVEAEDATYYCCQQMSSNPFFGGGTMLEI 105
||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db         62 SGSGSGTSYSLLTIKWMAEDATYYCCQQMSSNPLFFAGTKLEI 105
||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-08-211-202-3
; Sequence 3, Application US/08211202
; Patent No. 5565332
GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
APPLICANT: BAIER, Michael
APPLICANT: JESPERS, Laurent Stephane Anne Therese
APPLICANT: WINTER, Gregory Paul
TITLE OF INVENTION: Production of chimeric antibodies - a
TITLE OF INVENTION: combinatorial approach
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
```

```

ADDRESSSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
City: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120252.3
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120377.8
FILING DATE: 25-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEO ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-211-202-3

Query Match      87.4%: Score 486; DB 1; Length 107;
Best Local Similarity 88.6%: Pred. No. 9,1e-39;
Matches 93; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      2 IVLSGSPAILASPEGEKXTMTCRASSVNYMHWMYQKPGSSPKPMISATSNLASGVPARF 61
      1 |.|||||.....|
Db      2 IELTQSPILASPGKATMTTCRASSSVYMHWMYQKPGSSPKPMIVATSNLASGVPTRF 61
      1 |.|||||.....|

QY      62 SGSSGTSYSLTISRVEAEDATYYCQOMSSNPFTFGGTMLEIR 106
      1 |.|||||.....|
Db      62 SGTSGTSYSLTISRVEAEDATYYCQOMSSNPFTFGGTMLEIK 106
      1 |.|||||.....|

RESULT 15
US-08-469-486-57
: Sequence 57, Application US/08469486
: Patent No. 5739281
:
GENERAL INFORMATION:
:
APPLICANT: Theogersen, Hans Christian
:
APPLICANT: Hotzef, Thor Las
:
APPLICANT: Elzerodt, Michael
:
TITLE OF INVENTION: Improved method for the refolding of
:
TITLE OF INVENTION: proteins
:
NUMBER OF SEQUENCES: 58
:
CORRESPONDENCE ADDRESS:
:
ADDRESSEE: Fish & Richardson

```

```

101 STREET: 225 Franklin Street
102 CITY: Boston
103 STATE: Massachusetts
104 COUNTRY: USA
105 ZIP: 02110-2804
106
107 COMPUTER READABLE FORM:
108 MEDIUM TYPE: Floppy disk
109 COMPUTER: IBM PC compatible
110 OPERATING SYSTEM: PC-DOS/MS-DOS
111 SOFTWARE: PatentIn Release #1.0, Version
112 SOFTWARE: #1.25
113
114 CURRENT APPLICATION DATA:
115 APPLICATION NUMBER: US/08/469,486
116 FILING DATE:
117
118 CLASSIFICATION: 530
119
120 PRIOR APPLICATION DATA:
121 APPLICATION NUMBER: 08/192,060
122 FILING DATE: February 4, 1994
123 ATTORNEY/AGENT INFORMATION:
124 NAME: Paul T. Clark
125
126 REGISTRATION NUMBER: 30,162
127 REFERENCE/DOCKET NUMBER: 06363/002001
128
129 TELECOMMUNICATION INFORMATION:
130 TELEPHONE: 617 542 5070
131 TELEFAX: 617 542 8906
132 TELEX: 200154
133
134 INFORMATION FOR SEQ ID NO: 57:
135 SEQUENCE CHARACTERISTICS:
136 LENGTH: 246 amino acids
137 TYPE: amino acid
138 STRANDEDNESS:
139 TOPOLOGY: linear
140
141 MOLECULE TYPE: protein
142
143 US-08-469-486-57

```

[illegible]

Search completed: November 27, 2002, 07:25:14  
Job time : 21.2355 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:18:10 ; Search time 12.278 Seconds  
(without alignments)  
137.479 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 556  
Sequence: 1 QIVLSQSPAILASPEKVT.....COQSSNPPTFGCTMLEIR 106

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEM\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEM\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEM\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEM\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US10\_NEM\_PUB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US60\_NEM\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	100.0	106	US-09-893-615-89	Sequence 89, Appl
2	526	94.6	106	US-09-965-099-105	Sequence 105, App
3	526	94.6	106	US-10-051-852-105	Sequence 105, App
4	526	94.6	107	US-09-965-099-11	Sequence 11, Appl
5	526	94.6	107	US-10-051-852-11	Sequence 11, Appl
6	526	94.6	112	US-09-965-099-103	Sequence 103, App
7	526	94.6	112	US-10-051-852-103	Sequence 103, App
8	472	84.9	107	US-09-144-886-90	Sequence 90, Appl
9	464	83.5	109	US-09-144-886-91	Sequence 91, Appl
10	455	81.8	129	US-09-965-099-99	Sequence 99, Appl
11	455	81.8	129	US-10-051-852-99	Sequence 99, Appl
12	454	81.7	125	US-09-965-099-78	Sequence 78, Appl
13	454	81.7	125	US-10-051-852-78	Sequence 78, Appl
14	453	81.5	127	US-09-753-436-80	Sequence 80, Appl
15	451	81.1	107	US-09-965-099-62	Sequence 62, Appl
16	451	81.1	107	US-10-051-852-62	Sequence 62, Appl
17	449	80.8	107	US-09-144-886-82	Sequence 82, Appl
18	449	80.8	107	US-09-144-886-83	Sequence 83, Appl
19	448	80.6	106	US-09-976-787-24	Sequence 24, Appl

20	448	80.6	106	US-09-865-198-23	Sequence 23, Appl
21	448	80.6	108	US-09-976-787-8	Sequence 8, Appl
22	448	80.6	108	US-09-865-198-8	Sequence 8, Appl
23	448	80.6	238	US-09-976-787-29	Sequence 29, Appl
24	448	80.6	238	US-09-865-198-28	Sequence 28, Appl
25	448	80.6	240	US-09-976-787-28	Sequence 28, Appl
26	448	80.6	240	US-09-865-198-27	Sequence 27, Appl
27	447	80.4	107	US-09-982-107-12	Sequence 12, Appl
28	447	80.4	107	US-09-144-886-76	Sequence 76, Appl
29	447	80.4	107	US-09-965-099-57	Sequence 57, Appl
30	447	80.4	107	US-10-051-852-57	Sequence 57, Appl
31	446	80.2	109	US-09-965-099-95	Sequence 95, Appl
32	446	80.2	109	US-10-051-852-95	Sequence 95, Appl
33	445	80.0	235	US-09-910-059-17	Sequence 17, Appl
34	444	79.9	669	US-09-807-721-2	Sequence 2, Appl
35	443	79.7	108	US-09-910-059-9	Sequence 9, Appl
36	441	79.3	107	US-09-965-099-74	Sequence 74, Appl
37	441	79.3	107	US-10-051-852-74	Sequence 74, Appl
38	437	78.6	107	US-09-144-886-78	Sequence 78, Appl
39	437	78.6	119	US-09-808-037-28	Sequence 28, Appl
40	437	78.6	239	US-09-808-037-6	Sequence 6, Appl
41	434	78.1	107	US-09-144-886-88	Sequence 88, Appl
42	427	76.8	107	US-09-144-886-75	Sequence 75, Appl
43	422	75.9	109	US-09-144-886-97	Sequence 97, Appl
44	420	75.5	107	US-09-965-099-44	Sequence 44, Appl
45	420	75.5	107	US-10-051-852-44	Sequence 44, Appl

## ALIGNMENTS

RESULT 1  
US-09-893-615-89  
; Sequence 89, Application US/09893615  
; Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Filscher, Gerald W.  
Wong, Hing  
Schuman, Richard F.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSER: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 89:  
US-09-893-615-89

Query Match 100.0%; Score 556; DB 10; Length 106;  
Best Local Similarity 100.0%; Pred. No. 2,2e-33;  
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSOSPALISAPGKVTMTCRASSVNMHWYQKPGSSPKPWISATSNLASGVPAR 60  
DB 1 QIVLSOSPALISAPGKVTMTCRASSVNMHWYQKPGSSPKPWISATSNLASGVPAR 60  
QY 61 FSSGSGTSTSLTISRVEADAATYYCQWSSNPPTFGGTMLEIR 106  
DB 61 FSSGSGTSTSLTISRVEADAATYYCQWSSNPPTFGGTMLEIR 106

RESULT 2  
US-09-965-099-105  
Sequence 105, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Giora  
Patel, Arunbhai

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS

NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>

TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 105:

US-09-965-099-105

Query Match 94.6%; Score 526; DB 10; Length 106;  
Best Local Similarity 95.3%; Pred. No. 2,8e-31;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 QIVLSOSPALISAPGKVTMTCRASSVNMHWYQKPGSSPKPWISATSNLASGVPAR 60

DB 1 QIVLSOSPALISAPGKVTMTCRASSVNMHWYQKPGSSPKPWISATSNLASGVPAR 60

QY 61 FSSGSGTSTSLTISRVEADAATYYCQWSSNPPTFGGTMLEIR 106  
DB 61 FSSGSGTSTSLTISRVEADAATYYCQWSSNPPTFGGTMLEIR 106

RESULT 3  
US-10-051-852-105  
Sequence 105, Application US/10051852  
Patent No. US20020146411A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Giora  
Nichols, Andrew  
Padian, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS

NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>

TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 105:

US-10-051-852-105

Query Match 94.6%; Score 526; DB 12; Length 106;  
Best Local Similarity 95.3%; Pred. No. 2,8e-31;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 QIVLSOSPALISAPGKVTMTCRASSVNMHWYQKPGSSPKPWISATSNLASGVPAR 60

Db 1 QIVLSOSPAILISASPGKVTMTTCRASSSVNYMHVYQOKPGSSPKPWITATSNLASGVPAR 60  
QY 61 FSSGSGSTSYSLTISRVEADATYYCOQWSSNPPTFGGTMLEIR 106  
Db 61 FSSGSGSTSYSLTISRVEADATYYCOQWSSNPPTFGGTMLEIR 106

RESULT 4  
US-09-965-099-11

Sequence 11, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Giora  
Patel, Arunbhai  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-965-099-11

Query Match 94.6%; Score 526; DB 10; Length 107;  
Best Local Similarity 95.3%; Pred. No. 2,8e-31;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 QIVLSOSPAILISASPGKVTMTTCRASSSVNYMHVYQOKPGSSPKPWITATSNLASGVPAR 60  
Db 1 QIVLSOSPAILISASPGKVTMTTCRASSSVNYMHVYQOKPGSSPKPWITATSNLASGVPAR 60  
QY 61 FSSGSGSTSYSLTISRVEADATYYCOQWSSNPPTFGGTMLEIR 106  
Db 61 FSSGSGSTSYSLTISRVEADATYYCOQWSSNPPTFGGTMLEIR 106

RESULT 5

US-10-051-852-11  
Sequence 11, Application US/10051852  
Patent No. US20020146411A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Giora  
Nichols, Andrew  
Padlan, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-10-051-852-11

Query Match 94.6%; Score 526; DB 12; Length 107;  
Best Local Similarity 95.3%; Pred. No. 2,8e-31;  
Matches 101; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 QIVLSOSPAILISASPGKVTMTTCRASSSVNYMHVYQOKPGSSPKPWITATSNLASGVPAR 60  
Db 1 QIVLSOSPAILISASPGKVTMTTCRASSSVNYMHVYQOKPGSSPKPWITATSNLASGVPAR 60  
QY 61 FSSGSGSTSYSLTISRVEADATYYCOQWSSNPPTFGGTMLEIR 106  
Db 61 FSSGSGSTSYSLTISRVEADATYYCOQWSSNPPTFGGTMLEIR 106

RESULT 6  
US-09-965-099-103

Sequence 103, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Giora  
Patel, Arunbhai  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS,  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <unknown>  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
US-09-965-099-103

Query Match 94.6%, Score 526, DB 10, Length 112;  
Best Local Similarity 95.3%, Pred. No. 3e-31;  
Matches 101: Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Oy 1 QIVLSQSPALLISASPGKVTMTCRASSSVNWMYQOKPGSSPKPWISNTSLASGVAR 60  
Db 1 QIVLSQSPALLISASPGKVTMTCRASSSVNWMYQOKPGSSPKPWIVATSNLASGVAR 60  
Oy 61 FSGSGSGTSTSLTISRVEAEDATYYCCQWSSNPPTFGGTMLEIR 106  
Db 61 FSGSGSGTSTSLTISRVEAEDATYYCCQWSSNPPTFGGTMLEIR 106  
RESULT 7  
US-10-051-852-103  
Sequence 103, Application US/10051852  
Patent No. US20020146411A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Giora  
Nichols, Andrew

Padlan, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <unknown>  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
US-10-051-852-103

Query Match 94.6%, Score 526, DB 12, Length 112;  
Best Local Similarity 95.3%, Pred. No. 3e-31;  
Matches 101: Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Oy 1 QIVLSQSPALLISASPGKVTMTCRASSSVNWMYQOKPGSSPKPWISNTSLASGVAR 60  
Db 1 QIVLSQSPALLISASPGKVTMTCRASSSVNWMYQOKPGSSPKPWIVATSNLASGVAR 60  
Oy 61 FSGSGSGTSTSLTISRVEAEDATYYCCQWSSNPPTFGGTMLEIR 106  
Db 61 FSGSGSGTSTSLTISRVEAEDATYYCCQWSSNPPTFGGTMLEIR 106  
RESULT 8  
US-09-144-886-90  
Sequence 90, Application US/09144886  
Patent No. US2002015514A1  
GENERAL INFORMATION:  
APPLICANT: Marks, James D  
Amersdorfer, Peter  
TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize  
Botulinum Neurotoxins  
FILE REFERENCE: 2500 117USO  
CURRENT APPLICATION NUMBER: US/09/144,886

```
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 90
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: BONT/A clone
US-09-144-886-90

Query Match
Best Local Similarity 84.9%; Score 472; DB 9; Length 107;
Matches 89; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 4 LSQSALISASGPEKVTMTCRASSSVNHYQOKPGSSPKWISATSNLASGVPAR 63
DB 4 LTQSPALMSASGPEKVTMTCRASSSVSYWYQOKPGSSPRLMITYDTSNLASGV 63
OY 64 SSGSGSYSLTISRVEADATYCCQWSSNPPTFGGCTMLEIR 106
DB 64 SSGSGSYSLTISRVEADATYCCQWSSNPPTFGGCTMLEIR 106

RESULT 9
US-09-144-886-91
; Sequence 91, Application US/09144886
; Patent No. US2002015114A1
; GENERAL INFORMATION:
; APPLICANT: Marks, James D
; APPLICANT: Amersdorfer, Peter
; TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize
; TITLE OF INVENTION: Botulinum Neurotoxins
; FILE REFERENCE: 2500.117050
; CURRENT APPLICATION NUMBER: US/09/144.886
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 91
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: BONT/A clone
; OTHER INFORMATION: 3H4 region VL epitope 2
US-09-144-886-91

Query Match
Best Local Similarity 83.5%; Score 464; DB 9; Length 109;
Matches 90; Conservative 7; Mismatches 8; Indels 2; Gaps 1;

OY 2 IVLQSPALISASGPEKVTMTCRASSSV--NYMHYQOKPGSSPKWISATSNLASGV 59
DB 2 IELTQSPALMSASGPEKVTMTCRASSSVSYLVQWYQOKPGSSPRLMITYDTSNLASGV 61
OY 60 RFGSGSGSYSLTISRVEADATYCCQWSSNPPTFGGCTMLEIR 106
DB 62 RFGSGSGSYSLTISRVEADATYCCQWSSNPPTFGGCTMLEIR 108

RESULT 10
US-09-965-099-99
; Sequence 99, Application US/09965099
; Patent No. US20020136725A1
; GENERAL INFORMATION:
; APPLICANT: Blackburn, Michael
; APPLICANT: Feuerstein, Gloria
; APPLICANT: Patel, Arunbhai
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN
; TREATMENT OF THROMBOSIS
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESSES: 111
```

```
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/965,099
; FILING DATE: 26-Sep-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/346,487
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: P50438-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: <unknown>
; TELEX: <unknown>
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 99:
US-09-965-099-99

Query Match
Best Local Similarity 81.8%; Score 455; DB 10; Length 129;
Matches 84; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 1 QIVLSQPAIISASGPEKVTMTCRASSSVNHYQOKPGSSPKWISATSNLASGV 60
DB 20 QIVLTQSPALMSASGPEKVTMTCRASSSVNHYQOKPGKAPKPVYATSNLASGV 79
OY 61 FSGSGSGSYSLTISRVEADATYCCQWSSNPPTFGGCTMLEIR 106
DB 80 FSGSGSGSYSLTISRVEADATYCCQWSSNPPTFGGCTMLEIR 125

RESULT 11
US-10-051-852-99
; Sequence 99, Application US/10051852
; Patent No. US20020146411A1
; GENERAL INFORMATION:
; APPLICANT: Blackburn, Michael
; APPLICANT: Church, William
; APPLICANT: Gross, Mitchell
; APPLICANT: Feuerstein, Gloria
; APPLICANT: Nichols, Andrew
; APPLICANT: Padlan, Eduardo
; APPLICANT: Patel, Arunbhai
; APPLICANT: Sylvester, Daniel
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT
; OF THROMBOSIS
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
```

```

1 STATE: PA
2 COUNTRY: USA
3 ZIP: 19406
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Diskette
7 COMPUTER: IBM Compatible
8 OPERATING SYSTEM: DOS
9 SOFTWARE: FastSeq Version 1.5
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/10/051,852
13 FILING DATE: 17-Jan-2002
14 CLASSIFICATION: <unknown>
15
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: 09/344,050
18 FILING DATE: 25-JUN-1999
19 APPLICATION NUMBER: 08/783,853
20 FILING DATE: 16-JAN-1997
21
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Baumeister, Kirk
24 REGISTRATION NUMBER: 33,833
25 REFERENCE/DOCKET NUMBER: P50438
26
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 610-270-5096
29 TELEFAX: <unknown>
30
31 INFORMATION FOR SEQ ID NO: 99:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 129 amino acids
34 TYPE: amino acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37 MOLECULE TYPE: protein
38
39 HYPOTHETICAL: NO
40
41 ANTI-SENSE: NO
42
43 FRAGMENT TYPE: Internal
44
45 ORIGINAL SOURCE:
46 SEQUENCE DESCRIPTION: SEQ ID NO: 99:
47
48 US-10-051-852-99
49
50 Query Match 81.8% Score 455; DB 12; Length 129;
51 Best Local Similarity 79.2%; Pred. No. 3,3e-26;
52 Matches 84; Conservative 11; Mismatches 11; Indels 0; Gaps 0;
53
54 QY 1 QIVLSOSPALISASGEVYMTCRASSSVNWMYQOKPGSSPKPMISATSLASGVPAR 60
55 |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
56 20 QIVLTQSSSLASASGDRVTTTCRASSSVNWMYQOKPGKAPKPMIYATSLASGVPSR 79
57
58 QY 61 FSGSGSGTYSYLTISRVEADATYVCOQWSSNPTFGGTMLEIR 106
59 |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
60 80 FSGSGSGTDTLTISLQPEDPATYVCOQWSSINPTFGGTMLEIK 125
61
62 RESULT 12
63 US-09-965-099-78
64 Sequence 78, Application US/09965099
65 Patent No. US20020136725A1
66
67 GENERAL INFORMATION:
68 APPLICANT: Blackburn, Michael
69 Feuerstein, Gloria
70 Patel, Arunbhai
71
72 TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN
73 TREATMENT OF THROMBOSIS
74
75 NUMBER OF SEQUENCES: 111
76 CORRESPONDENCE ADDRESS:
77 ADDRESSEE: SmithKline Beecham Corporation
78 STREET: 709 Swedeland Road
79 CITY: King of Prussia
80 STATE: PA
81
82 COUNTRY: USA
83 ZIP: 19406
84
85 COMPUTER READABLE FORM:
86 MEDIUM TYPE: Diskette
87 COMPUTER: IBM Compatible
88

```

```

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ VERSION 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/965, 099
FILING DATE: 26-SEP-2001
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/346,487
ATTORNEY/AGENT INFORMATION:
NAME: Baumelster, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50438-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: <UNKNOWN>
TELEX: <UNKNOWN>
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 125 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 78:
US-09-965-099-78

Query Match      81.7%: Score 454; DB 10: Length 125;
Best Local Similarity 78.3%: Pred. NO. 3.8e-26;
Matches 83: Conservative 10; Mismatches 13; Indels 0; Gaps 0.

OY 1 QIVLSQSPAILASAPGEVYMTCTRASSSVNYVMHWYQOKRGSSPKPMISATSNLASGVPAR 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 20 QIVLQSPATLSLSRGEATLSRCRASSSVNYVMHWYQORGCAPKPMIVATSNLASGVPAR 79
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 FSGSGSGTYSILTISRVAEDAAITYCCQWSSNPRTFGGCTMLEIR 106
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 80 FSGSGSGTDTLTLSLEPEDFAVYYCQWSSINPRTFGGCTKVEIK 125
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-10-051-852-78
: Sequence 78, Application US/10051852
: Patent No. US2002014641A1
: GENERAL INFORMATION:
APPLICANT: Blackburn, Michael
      Church, William
      Gross, Mitchell
      Feuerstein, Gloria
      Michols, Andrew
      Padlan, Eduardo
      Patel, Arunbhai
      Sylvester, Daniel
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT
OF THROMBOSIS
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 125 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 78:  
US-10-051-852-78

Query Match 81.7%; Score 454; DB 12; Length 125;  
Best Local Similarity 78.3%; Pred. No. 3,8e-26;  
Matches 83; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Oy 1 QIVLSQSPAILASPGKVTMTCRASSSVNWMYQOKPGSSPKPWISATNSLASGVPAR 60  
Db 20 QIVLQSPAILASPGKVTMTCRASSSVNWMYQOKPGSPKMWIATNSLASGVPAR 79

Oy 61 FSGSGSGTSTLTISRVEADATYYCOOWSSNPPTFGGTMLEIR 106  
Db 80 FSGSGSGTSTLTISRVEADATYYCOOWSSNPPTFGGTMLEIR 125

RESULT 14  
US-09-753-436-80  
Sequence 80, Application US/09753436  
Patent No. US20010029293A1  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemay  
TITLE OF INVENTION: ICAM-Related Materials and Methods  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/753,436  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/382,289  
FILING DATE:  
APPLICATION NUMBER: US 08/487,113  
FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,754  
FILING DATE: 05-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Joseph A., Jr.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 33282  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 80:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-753-436-80

Query Match 81.5%; Score 453; DB 10; Length 127;  
Best Local Similarity 81.1%; Pred. No. 4.6e-26;  
Matches 86; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Oy 1 QIVLSQSPAILASPGKVTMTCRASSSVNWMYQOKPGSSPKPWISATNSLASGVPAR 60  
Db 22 QIVLQSPAILASPGKVTMTCRASSSVNWMYQOKPRSSPKPWIVLTSMLASGVPAR 81

Oy 61 FSGSGSGTSTLTISRVEADATYYCOOWSSNPPTFGGTMLEIR 106  
Db 82 FSGSGSGTSTLTISRVEADATYYCOOWKSIPITFGAGTMLEIR 127

RESULT 15  
US-09-965-099-62  
Sequence 62, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Gloria  
Patel, Arunbhal  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001

```
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/346,487
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50438-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: <Unknown>
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-09-965-099-62

Query Match      81.1%  Score 451;  DB 10;  Length 107;
Best Local Similarity 77.4%  Pred. No. 5.5e-26;
Matches 82;  Conservative 11;  Mismatches 13;  Indels 0;  Gaps 0;

QY  1 QIVLSQSPAILLSAPGEKVTMTGRASSSVNYMHMYOQKPGSSPKPWISATSNLASEGVPAR 60
    :|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB  1 EIVLTQSPATLSLSPGERATLSCRASSSVNYMHMYOQRPQCAPKPMIYATSNLASEGVPAR 60
    :|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY  61 FSGSGGTSTSLTISRYEADATYYCQOWSSNPTEGGGTMLEIR 106
    |||||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB  61 FSGSGGTDTLTITISLEPDAVYCCQOWSINPTFGGTKEIK 106
    |||||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

Search completed: November 27, 2002, 07:26:00  
Job time : 12.278 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:46:58 ; Search time 36.5677 Seconds  
(without alignments)  
3349.042 Million cell updates/sec

Title: US-09-893-615-88

Perfect score: 318  
Sequence: 1 CAATGTCCTCTCCAGTC.....GGACCATGCTGAATAAGA 318

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 341543 seqs, 192557720 residues

Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PC1\_NEM\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEM\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEM\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEM\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	318	100.0	318	10	US-09-893-615-88
2	300.4	94.5	321	10	US-09-965-099-6
3	293.6	93.0	318	10	US-10-051-852-6
4	293.6	93.0	318	10	US-09-965-099-104
5	293.6	93.0	318	12	US-10-051-852-104
6	293.6	93.0	335	12	US-09-965-099-102
7	293.6	93.0	335	12	US-10-051-852-102
8	268.4	84.4	390	10	US-09-753-436-79
9	268.4	83.1	357	10	US-09-808-037-27
10	268.4	83.1	717	10	US-09-808-037-5
11	261	82.1	318	10	US-09-976-787-77
12	261	82.1	318	10	US-09-865-198-26
13	261	82.1	324	10	US-09-976-787-16
14	261	82.1	324	10	US-09-865-198-16
15	261	82.1	714	10	US-09-976-787-31
16	261	82.1	720	10	US-09-976-787-30
17	259.6	81.6	322	9	US-09-982-107-11
18	257.8	81.1	357	10	US-09-910-059-8
19	257.8	81.1	705	10	US-09-910-059-16

20	251.4	79.1	717	8	US-08-940-544-3	Sequence 3, Appl1
21	251.4	79.1	2059	9	US-09-807-721-1	Sequence 1, Appl1
22	248.2	78.1	411	10	US-09-881-823-5	Sequence 5, Appl1
23	241.8	76.0	5691	10	US-09-897-006-11	Sequence 11, Appl1
24	198.6	62.5	321	10	US-09-965-099-56	Sequence 56, Appl1
25	198.6	62.5	321	12	US-10-051-852-96	Sequence 56, Appl1
26	198.6	62.3	412	10	US-09-965-099-98	Sequence 98, Appl1
27	198	62.3	412	12	US-10-051-852-98	Sequence 98, Appl1
28	197.8	62.2	357	10	US-09-753-436-95	Sequence 95, Appl1
29	195.4	61.4	321	10	US-09-965-099-61	Sequence 61, Appl1
30	195.4	61.4	321	12	US-10-051-852-61	Sequence 61, Appl1
31	195.2	61.4	330	10	US-09-965-099-94	Sequence 94, Appl1
32	195.2	61.4	330	12	US-10-051-852-94	Sequence 94, Appl1
33	194.8	61.3	401	10	US-09-965-099-77	Sequence 77, Appl1
34	194.8	61.3	401	12	US-10-051-852-77	Sequence 77, Appl1
35	193.8	60.9	321	10	US-09-965-099-43	Sequence 43, Appl1
36	193.8	60.9	321	12	US-10-051-852-43	Sequence 43, Appl1
37	169.8	53.4	280	10	US-09-965-099-34	Sequence 34, Appl1
38	169.8	53.4	280	12	US-10-051-852-34	Sequence 34, Appl1
39	166	52.2	321	10	US-09-965-099-73	Sequence 73, Appl1
40	166	52.2	321	12	US-10-051-852-73	Sequence 73, Appl1
41	156.4	49.2	321	10	US-09-910-059-64	Sequence 64, Appl1
42	156.4	49.2	705	10	US-09-910-059-96	Sequence 96, Appl1
43	154.6	48.6	321	10	US-09-910-059-70	Sequence 70, Appl1
44	154.6	48.6	705	10	US-09-910-059-98	Sequence 98, Appl1
45	152.2	47.9	321	10	US-09-910-059-60	Sequence 60, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-893-615-88  
Sequence 88, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Mong, Hing  
Schinson, Jeffrey L.  
Schuman, Richard F.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnadli, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04895, 0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..318  
SEQUENCE DESCRIPTION: SEQ ID NO: 88:  
US-09-893-615-88

Query Match 100.0%; Score 318; DB 10; Length 318;  
Best Local Similarity 100.0%; Pred. No. 1.7e-88;  
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATTTCTCTCCAGCTTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCA 60  
DB 1 CAAATTTCTCTCCAGCTTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCA 60  
QY 61 ATGACTTGGAGGGGCGAGCTCAAGTAAATTACATGCACTGGTACCGAGAGCCAGGA 120  
DB 61 ATGACTTGGAGGGGCGAGCTCAAGTAAATTACATGCACTGGTACCGAGAGCCAGGA 120  
QY 121 TCCCTCCCAACCCCTGATTTCTGCGACATCCAACTGGCTTCTGAGTCCCTGCTCGC 180  
DB 121 TCCCTCCCAACCCCTGATTTCTGCGACATCCAACTGGCTTCTGAGTCCCTGCTCGC 180  
QY 181 TTCAGTGGCAGTGGGTCTGGACCTTCTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240  
DB 181 TTCAGTGGCAGTGGGTCTGGACCTTCTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240  
QY 241 GATGCTGCCACTTATTACTGCCAGCACTGAGTAGTAACCCAGCCAGCTTCGGAGGGGG 300  
DB 241 GATGCTGCCACTTATTACTGCCAGCACTGAGTAGTAACCCAGCCAGCTTCGGAGGGGG 300  
QY 301 ACCATGCTGGAATAAGA 318  
DB 301 ACCATGCTGGAATAAGA 318

RESULT 2  
US-09-965-099-6  
Sequence 6, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:

APPLICANT: Blackburn, Michael  
Feuerstein, Gloria  
Patel, Arunbhai

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS

NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/965, 099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>

TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-965-099-6

Query Match 94.5%; Score 300.4; DB 10; Length 321;  
Best Local Similarity 96.5%; Pred. No. 4.3e-83;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATTTCTCTCCAGCTTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCA 60  
DB 1 CAAATTTCTCTCCAGCTTCCAGCAATCTCTGTCATCTCCAGGGGAAAGTCA 60  
QY 61 ATGACTTGGAGGGGCGAGCTCAAGTAAATTACATGCACTGGTACCGAGAGCCAGGA 120  
DB 61 ATGACTTGGAGGGGCGAGCTCAAGTAAATTACATGCACTGGTACCGAGAGCCAGGA 120  
QY 121 TCCCTCCCAACCCCTGATTTCTGCGACATCCAACTGGCTTCTGAGTCCCTGCTCGC 180  
DB 121 TCCCTCCCAACCCCTGATTTCTGCGACATCCAACTGGCTTCTGAGTCCCTGCTCGC 180  
QY 181 TTCAGTGGCAGTGGGTCTGGACCTTCTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240  
DB 181 TTCAGTGGCAGTGGGTCTGGACCTTCTACTCTCTCACAATCAGCAGAGTGAAGCTGAA 240  
QY 241 GATGCTGCCACTTATTACTGCCAGCACTGAGTAGTAACCCAGCCAGCTTCGGAGGGGG 300  
DB 241 GATGCTGCCACTTATTACTGCCAGCACTGAGTAGTAACCCAGCCAGCTTCGGAGGGGG 300  
QY 301 ACCATGCTGGAATAAGA 318  
DB 301 ACCATGCTGGAATAAGA 318

RESULT 3  
US-10-051-852-6  
Sequence 6, Application US/10051852  
Patent No. US20020146411A1  
GENERAL INFORMATION:

APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Gloria  
Nichols, Andrew  
Padlan, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS

NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-051-852-6

Query Match 94.5%; Score 300.4; DB 12; Length 321;  
Best Local Similarity 96.5%; Pred. No. 4.3e-83;  
Matches 307; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAAGTCCACA 60  
DB 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAAGTCCACA 60  
QY 61 ATGACTTGGAGGGCCAGCTCAAGTGAATTAATGACATGCTGATCCAGCAGAAAGCCAGGA 120  
DB 61 ATGACTTGGAGGGCCAGCTCAAGTGAATTAATGACATGCTGATCCAGCAGAAAGCCAGGA 120  
QY 121 TCCGCCCAAAACCTGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180  
DB 121 TCCGCCCAAAACCTGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180  
QY 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAA 240  
DB 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAA 240  
QY 241 GATGCTGCCACTTATTAAGCCGACAGTGGAGTAAACCCACCGTTCGGAGGGGG 300  
DB 241 GATGCTGCCACTTATTAAGCCGACAGTGGAGTAAACCCACCGTTCGGAGGGGG 300  
QY 301 ACCACTGTGAAATAGA 318  
DB 301 ACCACTGTGAAATAGA 318

RESULT 4  
US-09-965-099-104  
Sequence 104, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Gloria  
Patel, Arunbhai  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESS: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road

CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1..318  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 104:  
US-09-965-099-104

Query Match 93.0%; Score 295.6; DB 10; Length 318;  
Best Local Similarity 95.6%; Pred. No. 1.3e-81;  
Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAAGTCCACA 60  
DB 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGATCTCCAGGGGAAAAGTCCACA 60  
QY 61 ATGACTTGGAGGGCCAGCTCAAGTGAATTAATGACATGCTGATCCAGCAGAAAGCCAGGA 120  
DB 61 ATGACTTGGAGGGCCAGCTCAAGTGAATTAATGACATGCTGATCCAGCAGAAAGCCAGGA 120  
QY 121 TCCGCCCAAAACCTGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180  
DB 121 TCCGCCCAAAACCTGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180  
QY 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAA 240  
DB 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAA 240  
QY 241 GATGCTGCCACTTATTAAGCCGACAGTGGAGTAAACCCACCGTTCGGAGGGGG 300  
DB 241 GATGCTGCCACTTATTAAGCCGACAGTGGAGTAAACCCACCGTTCGGAGGGGG 300  
QY 301 ACCACTGTGAAATAGA 318  
DB 301 ACCACTGTGAAATAGA 318

RESULT 5  
US-10-051-852-104

Sequence 104, Application US/10051852  
Patent No. US20020146411A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Giora  
Nichols, Andrew  
Padlan, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...318  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 104:  
US-10-051-852-104  
Query Match 93.0%; Score 295.6; DB 12; Length 318;  
Best Local Similarity 95.6%; Pred. No. 1.3e-81;  
Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CAATTTGTCCTCCAGCTCCAGCAATCCTGTCGATCTCCAGGGGAAAGGTGACA 60  
D 1 CAGATAGACTCTCCAGCTCCAGCAATCCTGTCGATCTCCAGGGGAAAGGTGACA 60  
QY 61 ATGACTGACAGGCGACGTCAAGTGAATTAATGACGCTGTAACGACAGAGCCAGAGA 120  
D 61 ATGACTGACAGGCGACGTCAAGTGAATTAATGACGCTGTAACGACAGAGCCAGAGA 120

QY 121 TCCCTCCCAAAACCCGTGATTTTCGCACATCCAAACCTGGCTTCTGCAGTCCCTGCTGC 180  
D 121 TCCCTCCCAAAACCCGTGATTTATGCACATCCAAACCTGGCTTCTGCAGTCCCTGCTGC 180  
QY 181 TTCAGTGGCAATGGGTCTGGGACCTCTTACTCTCAATCAATCAGCAATGTGGAGGTGAA 240  
D 181 TTCAGTGGCAATGGGTCTGGGACCTCTTACTCTCAATCAATCAGCAATGTGGAGGTGAA 240  
QY 241 GATCTGCCACTTTATTACTGCCAGAGTAGTAGTAACCCAGCCAGCTTGGAGGGGG 300  
D 241 GATCTGCCACTTTATTACTGCCAGAGTAGTAGTAACCCAGCCAGCTTGGAGGGGG 300  
QY 301 ACCAAGCTGGAATCAGA 318  
D 301 ACCAAGCTGGAATCAGA 318  
RESULT 6  
US-09-965-099-102  
Sequence 102, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Giora  
Patel, Arunbhai  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 335 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...335  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 102:  
US-09-965-099-102

Query Match 93.0%; Score 295.6; DB 10; Length 335;  
Best Local Similarity 95.6%; Pred. No. 1.3e-81;  
Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CAAATGTTCTCTCCAGCTCTCCAGCAATCTGTGCAATCCAGGGGAAAGGTGACA 60  
1 CAGATAGTACTCTCCAGCTCTCCAGCAATCTGTGCAATCCAGGGGAAAGGTGACA 60  
Db 1 CAGATAGTACTCTCCAGCTCTCCAGCAATCTGTGCAATCCAGGGGAAAGGTGACA 60  
Qy 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTCATGACATGCTACCAGCAGAAAGCAGGA 120  
Db 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTCATGACATGCTACCAGCAGAAAGCAGGA 120  
Qy 121 TCCTCCCCCAAAACCTGATTTCTCCACATCCAGCTGGCTTGGAGTCCCTGCTGCG 180  
Db 121 TCCTCCCCCAAAACCTGATTTATGCAATCCAGCTGGCTTGGAGTCCCTGCTGCG 180  
Qy 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGAGAGAGGCTGAA 240  
Db 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGAGAGAGGCTGAA 240  
Qy 241 GATGCTGCACCTATTACTGCGCAGCAGTGGAGTAAACCCACGCTTGGAGGGGG 300  
Db 241 GATGCTGCACCTATTACTGCGCAGCAGTGGAGTAAACCCACGCTTGGAGGGGG 300  
Qy 301 ACCATGCTGGAATAGA 318  
Db 301 ACCAAGCTGGAATCAAA 318

RESULT 7  
US-10-051-852-102  
Sequence 102, Application US/10051852  
Patent No. US20020146411A1

## GENERAL INFORMATION:

APPLICANT: Blackburn, Michael

Church, William

Gross, Mitchell

Feuerstein, Gloria

Nichols, Andrew

Padlan, Eduardo

Patel, Arunbhai

Sylvester, Daniel

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT OF THROMBOSIS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/051,852

FILING DATE: 17-Jan-2002

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/344,050

FILING DATE: 25-JUN-1999

APPLICATION NUMBER: 08/783,853

FILING DATE: 16-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Baumelster, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: P50438

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096

TELEFAX: &lt;Unknown&gt;

TELEX: <Unknown>  
INFORMATION FOR SEQ. ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 335 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding sequence  
LOCATION: 1...335  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 102:  
US-10-051-852-102

Query Match 93.0%; Score 295.6; DB 12; Length 335;  
Best Local Similarity 95.6%; Pred. No. 1.3e-81;  
Matches 304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CAAATGTTCTCTCCAGCTCTCCAGCAATCTGTGCAATCCAGGGGAAAGGTGACA 60  
1 CAGATAGTACTCTCCAGCTCTCCAGCAATCTGTGCAATCCAGGGGAAAGGTGACA 60  
Db 1 CAGATAGTACTCTCCAGCTCTCCAGCAATCTGTGCAATCCAGGGGAAAGGTGACA 60  
Qy 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTCATGACATGCTACCAGCAGAAAGCAGGA 120  
Db 61 ATGACTTGACAGGGCCAGCTCAAGTAAATTCATGACATGCTACCAGCAGAAAGCAGGA 120  
Qy 121 TCCTCCCCCAAAACCTGATTTCTCCACATCCAGCTGGCTTGGAGTCCCTGCTGCG 180  
Db 121 TCCTCCCCCAAAACCTGATTTATGCAATCCAGCTGGCTTGGAGTCCCTGCTGCG 180  
Qy 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGAGAGAGGCTGAA 240  
Db 181 TTCAGTGGCAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGAGAGAGGCTGAA 240  
Qy 241 GATGCTGCACCTATTACTGCGCAGCAGTGGAGTAAACCCACGCTTGGAGGGGG 300  
Db 241 GATGCTGCACCTATTACTGCGCAGCAGTGGAGTAAACCCACGCTTGGAGGGGG 300  
Qy 301 ACCATGCTGGAATAGA 318  
Db 301 ACCAAGCTGGAATCAAA 318

RESULT 8  
US-09-753-436-79  
Sequence 79, Application US/09753436  
Patent No. US20010029293A1

## GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael

Vazeux, Rosemary

TITLE OF INVENTION: ICAI-Related Materials and Methods

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &amp; Borum

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/753,436

FILING DATE:

CLASSIFICATION:

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/382,289
; FILING DATE:
; APPLICATION NUMBER: US 08/487,113
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,754
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Joseph A., Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 33282
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-09-753-436-79

Query Match      84.4%; Score 268.4; DB 10; Length 390;
Best Local Similarity 90.3%; Pred. No. 3.1e-73;
Matches 287; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCTGTGTCATCTCCAGGGGAAAGGTACA 60
DB 73 CAAATGTTCTCAACCGAGTCTCCAGCACTATGTCATCTCCAGGGGAAAGGTACC 132
QY 61 ATGACTTGACAGGGGCACTCAAGTGAATTACATGCACTGCTGTAACGACGAGAGCCAGGA 120
DB 133 ATGACCTGACAGTGCACACTCAAGTGAAGTTACATTTATTGTTACGACGAGAGCCAGGA 192
QY 121 TCCGCCCCCAACCGCTGATTTCTGCCACATCCAACTGGCTTGTGAGTCCCGCTCGCG 180
DB 193 TCCGCCCCCAACCGCTGATTTCTGCCATCTCAATCCAACTGGCTTGTGAGTCCCGCTCGCG 252
QY 181 TTCAGTGGCAGTGGGCTTGGAGACTTCTTACTCTCTCAATTCACAGACAGTGGAGGCTGAA 240
DB 253 TTCAGTGGCAGTGGGCTTGGAGGCTTCTTACTCTCTCAATTCACAGACAGTGGAGGCTGAA 312
QY 241 GATGCTCCCACTTATTAATCTCCAGCACTGAGTGAATACCCACCCAGCTTGGAGGGGGG 300
DB 313 GATGCTCCCACTTATTAATCTCCAGCACTGAGTGAATACCCACCTACCTGCTGGTGGG 372
QY 301 ACCATGCTGGAATTAAGA 318
DB 373 ACCAAGCTGGAGCTGAAGA 390

RESULT 9
US-09-808-037-27
; Sequence 27, Application US/09808037
; Patent No. US20020052311A1
```

```

; GENERAL INFORMATION:
; APPLICANT: SOLOMON, Bekka
; APPLICANT: HANAN, Eliat
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF
; TITLE OF INVENTION: NEUROLOGICAL DISEASES AND DISORDERS
; FILE REFERENCE: SOLOMON-2D
; CURRENT APPLICATION NUMBER: US/09/808,037
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/629,971
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 09/473,653
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 60/152,417
; PRIOR FILING DATE: 1999-09-03
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 357
; TYPE: DNA
; ORGANISM: synthetic construct
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(357)
; US-09-808-037-27

Query Match      83.1%; Score 264.2; DB 10; Length 357;
Best Local Similarity 89.6%; Pred. No. 5.8e-72;
Matches 284; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2 AAATGTTCTCTCCAGTCTCCAGCAATCTGTGTCATCTCCAGGGGAAAGGTACAA 61
DB 41 ACATGAGCTCACTCACTCTCCAGCAATCTGTGTCATCTCCAGGGGAAAGGTACCA 100
QY 62 TGAATTCAGAGGCGAGTCAAGTGAATTAATCATGATGATGATACAGACAGAGGAGAT 121
DB 101 TGACCTGCACTGAGTCAAGTGAATTAATCATGATGATGATACAGACAGAGGAGAT 160
QY 122 CTTCCGCCCAACCGCTGATTTCTGCCACATCCAACTGGCTTGTGAGTCCCGCTGCT 181
DB 161 CTTCCGCCCAACCGATGATTTATGACATCCAACTGGCTTGTGAGTCCCGCTGCT 220
QY 182 TCAATGTCAGTGGTCTGGAGACCTTACTCTCTCAATTCACAGACAGTGGAGGCTGAG 241
DB 221 TCAATGTCAGTGGTCTGGAGACCTTACTCTCTCAATTCACAGACAGTGGAGGCTGAG 280
QY 242 ATGCTGCACTTATTAATCTCCAGCACTGAGTGAATACCCACCCAGCTTGGAGGGGGA 301
DB 281 ATGCTGCACTTATTAATCTCCAGCACTGAGTGAATACCCACCTGCTGGAGGGGGA 340
QY 302 CCATGCTGGAATTAAGA 318
DB 341 CCAAGCTGGAATTAAGA 357

RESULT 10
US-09-808-037-5
; Sequence 5, Application US/09808037
; Patent No. US20020052311A1
; GENERAL INFORMATION:
; APPLICANT: SOLOMON, Bekka
; APPLICANT: HANAN, Eliat
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF
; TITLE OF INVENTION: NEUROLOGICAL DISEASES AND DISORDERS
; FILE REFERENCE: SOLOMON-2D
; CURRENT APPLICATION NUMBER: US/09/808,037
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/629,971
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 09/473,653
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 60/152,417
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 33
```



```
APPLICANT: Zhu, Zhenping
APPLICANT: Wille, Larry
TITLE OF INVENTION: Antibodies Specific to KDR and Uses Thereof
FILE REFERENCE: 11245/46505
CURRENT APPLICATION NUMBER: US/09/976,787
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 09/493,539
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,726
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 40
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 16
LENGTH: 324
TYPE: DNA
ORGANISM: Mus musculus
US-09-976-787-16

Query Match
Best Local Similarity 82.1%; Score 261; DB 10; Length 324;
Matches 282; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 2 AAATTGTTCTCTCCAGCTCCAGCAATCCTGTGATCTCCAGGGGAAAGGTACAA 61
DB 2 ACATCGAGCTCAGCTCAGCTCCAGCAATCCTGTGATCTCCAGGGGAGAGGTACAA 61
QY 62 TGACTTGACAGGGCCAGCTCAAGTGTAAATTCATGCACTGGTACCAAGAGGCTAG 121
DB 62 TAACCTGCACTGCGAGCTCAAGTGTAAATTCATGCACTGGTACCAAGAGGCTAG 121
QY 122 CTTCTCCCAAACTGTGATTTATAGCAATCCAACTGGCTTCGGAGTCCCTGCTGCT 181
DB 122 CTTCTCCCAAACTGTGATTTATAGCAATCCAACTGGCTTCGGAGTCCCTGCTGCT 181
QY 182 TCAGTGGCAGTGGGTGGGACCTTACTCTGTCCACATCAGCAGGAGGCTGAG 241
DB 182 TCAGTGGCAGTGGGTGGGACCTTACTCTGTCCACATCAGCAGGAGGCTGAG 241
QY 242 ATGCTGCACTTATTACTGCGCAGAGTAGTAGTAACCCAGCCAGCTTGGAGGGGGA 301
DB 242 ATGCTGCACTTATTACTGCGCAGAGTAGTAGTAACCCAGCTTGGAGGGGGA 301
QY 302 CCATGCTGGAATTAAGA 318
DB 302 CCATGCTGGAATTAAGA 318
DB 302 CCAAGCTGGAATTAAGA 318
```

```
RESULT 14
US-09-865-198-16
Sequence 16, Application US/09865198
Patent No. US20020103345A1
GENERAL INFORMATION:
APPLICANT: Zhu, Zhenping
TITLE OF INVENTION: Bispecific Immunoglobulin-Like Antigen Binding Proteins and Methods
FILE REFERENCE: 11245/47102
CURRENT APPLICATION NUMBER: US/09/865,198
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 60/206,749
PRIOR FILING DATE: 2000-05-24
NUMBER OF SEQ ID NOS: 34
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 16
LENGTH: 324
TYPE: DNA
ORGANISM: Mouse
US-09-865-198-16
```

```
Query Match
Best Local Similarity 82.1%; Score 261; DB 10; Length 324;
Matches 282; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 2 AAATTGTTCTCTCCAGCTCCAGCAATCCTGTGATCTCCAGGGGAAAGGTACAA 61
```

```
DB 2 ACATCGAGCTCAGCTCAGCTCCAGCAATCCTGTGATCTCCAGGGGAGAGGTACAA 61
QY 62 TGACTTGACAGGGCCAGCTCAAGTGTAAATTCATGCACTGGTACCAAGAGGCTAG 121
DB 62 TAACCTGCACTGCGAGCTCAAGTGTAAATTCATGCACTGGTACCAAGAGGCTAG 121
QY 122 CTTCTCCCAAACTGTGATTTATAGCAATCCAACTGGCTTCGGAGTCCCTGCTGCT 181
DB 122 CTTCTCCCAAACTGTGATTTATAGCAATCCAACTGGCTTCGGAGTCCCTGCTGCT 181
QY 182 TCAGTGGCAGTGGGTGGGACCTTACTCTGTCCACATCAGCAGGAGGCTGAG 241
DB 182 TCAGTGGCAGTGGGTGGGACCTTACTCTGTCCACATCAGCAGGAGGCTGAG 241
QY 242 ATGCTGCACTTATTACTGCGCAGAGTAGTAGTAACCCAGCCAGCTTGGAGGGGGA 301
DB 242 ATGCTGCACTTATTACTGCGCAGAGTAGTAGTAACCCAGCTTGGAGGGGGA 301
QY 302 CCATGCTGGAATTAAGA 318
DB 302 CCATGCTGGAATTAAGA 318
DB 302 CCAAGCTGGAATTAAGA 318
```

```
RESULT 15
US-09-976-787-31
Sequence 31, Application US/09976787
Patent No. US20020064528A1
GENERAL INFORMATION:
APPLICANT: Zhu, Zhenping
APPLICANT: Wille, Larry
TITLE OF INVENTION: Antibodies Specific to KDR and Uses Thereof
FILE REFERENCE: 11245/46505
CURRENT APPLICATION NUMBER: US/09/976,787
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 09/493,539
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,726
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 40
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 31
LENGTH: 714
TYPE: DNA
ORGANISM: Mouse
US-09-976-787-31
```

```
Query Match
Best Local Similarity 82.1%; Score 261; DB 10; Length 714;
Matches 282; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 2 AAATTGTTCTCTCCAGCTCCAGCAATCCTGTGATCTCCAGGGGAAAGGTACAA 61
DB 398 ACATCGAGCTCAGCTCAGCTCCAGCAATCCTGTGATCTCCAGGGGAGAGGTACAA 457
QY 62 TGACTTGACAGGGCCAGCTCAAGTGTAAATTCATGCACTGGTACCAAGAGGCTAG 121
DB 62 TAACCTGCACTGCGAGCTCAAGTGTAAATTCATGCACTGGTACCAAGAGGCTAG 121
QY 122 CTTCTCCCAAACTGTGATTTATAGCAATCCAACTGGCTTCGGAGTCCCTGCTGCT 181
DB 122 CTTCTCCCAAACTGTGATTTATAGCAATCCAACTGGCTTCGGAGTCCCTGCTGCT 181
QY 182 TCAGTGGCAGTGGGTGGGACCTTACTCTGTCCACATCAGCAGGAGGCTGAG 241
DB 182 TCAGTGGCAGTGGGTGGGACCTTACTCTGTCCACATCAGCAGGAGGCTGAG 241
QY 242 ATGCTGCACTTATTACTGCGCAGAGTAGTAGTAACCCAGCCAGCTTGGAGGGGGA 301
DB 242 ATGCTGCACTTATTACTGCGCAGAGTAGTAGTAACCCAGCTTGGAGGGGGA 301
QY 302 CCATGCTGGAATTAAGA 318
DB 302 CCATGCTGGAATTAAGA 318
```



Wed Nov 27 08:54:28 2002

us-09-893-615-88.rnpb

Page 9

Db 698 CCAGCTGGGAATAAAA 714

Search completed: November 27, 2002, 05:31:02  
Job time : 38.5677 secs

---

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:18:10 : Search time 14.2471 Seconds  
(without alignments)  
137.479 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 1 EVMALVSGGGLVQPKGSLKL.....SGIDYANDYWGQGTSLTVSS 123

Sequence: 1 EVMALVSGGGLVQPKGSLKL.....SGIDYANDYWGQGTSLTVSS 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCRT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubppaa/PCRTUS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	646	100.0	123	10	US-09-893-615-87
2	532.5	82.4	143	10	US-09-881-823-16
3	519	80.3	117	10	US-09-835-087-8
4	519	80.3	117	10	US-09-809-739-12
5	519	80.3	117	10	US-09-840-459-10
6	519	80.3	148	10	US-09-840-459-100
7	493	76.3	117	10	US-09-835-087-13
8	493	76.3	117	10	US-09-809-739-23
9	493	76.3	117	10	US-09-840-459-20
10	489	75.7	117	10	US-09-835-087-12
11	489	75.7	117	10	US-09-809-739-22
12	489	75.7	117	10	US-09-840-459-19
13	488	75.5	117	10	US-09-835-087-11
14	488	75.5	117	10	US-09-809-739-21
15	488	75.5	117	10	US-09-840-459-18
16	487	75.4	117	10	US-09-835-087-10
17	487	75.4	117	10	US-09-809-739-20
18	487	75.4	117	10	US-09-840-459-17
19	487	75.4	119	10	US-09-840-459-104

20	483	74.8	101	10	US-09-840-459-34	Sequence 34, Appl
21	471.5	73.0	126	10	US-09-840-459-74	Sequence 74, Appl
22	468	72.4	100	10	US-09-840-459-35	Sequence 35, Appl
23	463	71.7	127	10	US-09-840-459-71	Sequence 71, Appl
24	450.5	71.3	120	12	US-10-025-687-4	Sequence 4, Appl
25	455.5	70.5	126	10	US-09-840-459-73	Sequence 73, Appl
26	444.5	68.8	263	9	US-09-956-086-3	Sequence 3, Appl
27	444.5	68.8	263	9	US-09-956-087-3	Sequence 6, Appl
28	444.5	68.8	283	9	US-09-983-442-6	Sequence 8, Appl
29	444.5	68.8	283	10	US-09-983-580-6	Sequence 8, Appl
30	443	68.4	123	10	US-09-840-459-82	Sequence 4, Appl
31	442	68.4	140	10	US-09-286-240-4	Sequence 72, Appl
32	438.5	67.9	126	10	US-09-840-459-72	Sequence 91, Appl
33	438	67.8	123	10	US-09-840-459-91	Sequence 89, Appl
34	436.5	67.6	124	10	US-09-840-459-89	Sequence 84, Appl
35	435	67.3	125	10	US-09-840-459-84	Sequence 2, Appl
36	434.5	67.3	298	10	US-09-883-758-2	Sequence 3, Appl
37	434	67.2	119	10	US-09-811-123-3	Sequence 28, Appl
38	434	67.2	121	10	US-09-840-459-92	Sequence 90, Appl
39	433.5	67.1	120	10	US-09-229-200A-28	Sequence 131, App
40	433.5	67.1	124	10	US-09-840-459-90	Sequence 5, Appl
41	433	67.0	473	10	US-09-910-059-131	Sequence 5, Appl
42	432.5	67.0	119	10	US-09-756-301A-5	Sequence 5, Appl
43	432.5	67.0	119	10	US-09-927-703-5	Sequence 5, Appl
44	432.5	67.0	119	10	US-09-766-535A-5	Sequence 5, Appl
45	432.5	67.0	119	10	US-09-756-161A-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-893-615-87  
Sequence 87, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR POSITIVE BACTERIA  
NUMBER OF SEQUENCE ADDRESSES: 89  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995, 0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 123 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 87;  
US-09-893-615-87

Query Match 100.0%; Score 646; DB 10; Length 123;  
Best Local Similarity 100.0%; Pred. No. 1.5e-52;  
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVMALVESGGGLVQPKGSLKLSCAASGFTFNNYAMNVRQAPGKGLEWVARIRSKSNYYAT 60  
|||||  
DB 1 EVMALVESGGGLVQPKGSLKLSCAASGFTFNNYAMNVRQAPGKGLEWVARIRSKSNYYAT 60  
QY 61 FYADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVRGASGIDYANDYWGQGTSLT 120  
|||||  
DB 61 FYADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVRGASGIDYANDYWGQGTSLT 120  
QY 121 VSS 123  
|||  
DB 121 VSS 123

RESULT 2  
US-09-881-823-16

Sequence 16, Application US/09881823  
Patent No. US2002006806A1  
GENERAL INFORMATION:  
APPLICANT: SHI, WENYUAN  
APPLICANT: ANDERSON, MAXWELL  
APPLICANT: MORRISON, SHERIE  
APPLICANT: TRINH, RYAN  
APPLICANT: WIMS, LETITIA  
APPLICANT: CHEN, LI  
TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries  
FILE REFERENCE: 22851-032  
CURRENT APPLICATION NUMBER: US/09/881,823  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 07/378,577  
PRIOR FILING DATE: 1999-08-20  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 16  
LENGTH: 143  
TYPE: PRT  
ORGANISM: Murine  
US-09-881-823-16

Query Match 82.4%; Score 532.5; DB 10; Length 143;  
Best Local Similarity 84.8%; Pred. No. 3.9e-42;  
Matches 106; Conservative 5; Mismatches 9; Indels 5; Gaps 2;

QY 1 EVMALVESGGGLVQPKGSLKLSCAASGFTFNNYAMNVRQAPGKGLEWVARIRSKSNYYAT 60  
|||||  
DB 20 EVOLVETGGGLVQPKGSLKLSCAASGFTFNTAMNVRQAPGKGLEWVARIRSKSNYYAT 79  
QY 61 FYADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVRGASGIDYA--MDYWGQGT 118  
|||||  
DB 80 YVADSVEDRTISRDSQSMLYLQMNMLKTEDTAMYYCVR---NYDYDAMSAYWGGTV 136  
QY 119 LTVSS 123  
|||||  
DB 137 VTVSS 141

RESULT 3  
US-09-835-087-8

Sequence 8, Application US/09835087  
Patent No. US20020042370A1  
GENERAL INFORMATION:  
APPLICANT: Wayne W. Hancock  
TITLE OF INVENTION: Method of Treating Graft Rejection Using  
FILE REFERENCE: 1855.2008-003  
CURRENT APPLICATION NUMBER: US/09/835,087  
CURRENT FILING DATE: 2001-09-24

PRIOR APPLICATION NUMBER: 09/549,448  
PRIOR FILING DATE: 2000-04-14  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 117  
TYPE: PRT  
ORGANISM: Mus Musculus  
FEATURE:  
NAME/KEY: DOKAIN  
LOCATION: (1)...(117)

US-09-835-087-8  
OTHER INFORMATION: Murine mAb 1D9 heavy chain variable region

Query Match 80.3%; Score 519; DB 10; Length 117;  
Best Local Similarity 80.5%; Pred. No. 5.4e-41;  
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;

QY 1 EVMALVESGGGLVQPKGSLKLSCAASGFTFNNYAMNVRQAPGKGLEWVARIRSKSNYYAT 60  
|||||  
DB 1 EVOLVESGGGLVQPKGSLKLSCAASGFTFNNYAMNVRQAPGKGLEWVARIRSKSNYYAT 60  
QY 61 FYADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVRGASGIDYANDYWGQGTSLT 120  
|||||  
DB 61 YVADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVRGASGIDYANDYWGQGTSLT 114  
QY 121 VSS 123  
|||  
DB 115 VSS 117

RESULT 4  
US-09-809-739-12

Sequence 12, Application US/09809739  
Patent No. US20020106369A1  
GENERAL INFORMATION:  
APPLICANT: Horvath, Christopher J.  
APPLICANT: Rao, Patricia E.  
TITLE OF INVENTION: Method of Inhibiting Stenosis and  
FILE REFERENCE: 1855.1069-003  
CURRENT APPLICATION NUMBER: US/09/809,739  
CURRENT FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: US 09/528,267  
PRIOR FILING DATE: 2000-03-17  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 117  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1)...(117)  
OTHER INFORMATION: Murine mAb 1D9 heavy chain variable region  
NAME/KEY: SITE  
LOCATION: (31)...(35)  
OTHER INFORMATION: CDR1  
NAME/KEY: SITE  
LOCATION: (50)...(68)  
OTHER INFORMATION: CDR2  
NAME/KEY: SITE  
LOCATION: (101)...(106)  
OTHER INFORMATION: CDR3  
OTHER INFORMATION: Mouse

US-09-809-739-12

Query Match 80.3%; Score 519; DB 10; Length 117;  
Best Local Similarity 80.5%; Pred. No. 5.4e-41;  
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;

QY 1 EVMALVESGGGLVQPKGSLKLSCAASGFTFNNYAMNVRQAPGKGLEWVARIRSKSNYYAT 60  
|||||

```
Db 1 EVOLVESGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRTKNNVAT 60
QY 61 FYADSVKDRFTTSRDSSQSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYGQTSILT 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YYADSVKDRYTTISRDSSEMLFLQNNLKTEDTAMYYCVTFYNGV-----WGTGTVT 114
QY 121 VSS 123
|||
Db 115 VSS 117

RESULT 5
US-09-840-459-10
; Sequence 10, Application US/09840459
; Patent No. US20020150576A1
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-840-459-10

Query Match 80.3%; Score 519; DB 10; Length 117;
Best Local Similarity 80.5%; Pred. No. 5,4e-41;
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;

QY 1 EYWLVSGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRKSNVAT 60
|||
Db 1 EVOLVESGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRTKNNVAT 60
QY 61 FYADSVKDRFTTSRDSSQSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYGQTSILT 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YYADSVKDRYTTISRDSSEMLFLQNNLKTEDTAMYYCVTFYNGV-----WGTGTVT 114
QY 121 VSS 123
|||
Db 115 VSS 117

RESULT 6
US-09-840-459-100
; Sequence 100, Application US/09840459
; Patent No. US20020150576A1
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 100
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-840-459-100

Query Match 80.3%; Score 519; DB 10; Length 148;
Best Local Similarity 80.5%; Pred. No. 6,9e-41;
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;

QY 1 EYWLVSGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRKSNVAT 60
|||
Db 20 EVOLVESGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRTKNNVAT 79
QY 61 FYADSVKDRFTTSRDSSQSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYGQTSILT 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 80 YYADSVKDRYTTISRDSSEMLFLQNNLKTEDTAMYYCVTFYNGV-----WGTGTVT 133
QY 121 VSS 123
|||
Db 134 VSS 136

RESULT 7
US-09-835-087-13
; Sequence 13, Application US/09835087
; Patent No. US20020042370A1
; GENERAL INFORMATION:
; APPLICANT: Wayne W. Hancock
; TITLE OF INVENTION: Method of Treating Graft Rejection Using
; FILE REFERENCE: 1855.2008-003
; CURRENT APPLICATION NUMBER: US/09/835,087
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/549,448
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: Humanized sequence
US-09-835-087-13

Query Match 76.3%; Score 493; DB 10; Length 117;
Best Local Similarity 76.4%; Pred. No. 1,3e-38;
Matches 94; Conservative 14; Mismatches 9; Indels 6; Gaps 1;

QY 1 EYWLVSGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRKSNVAT 60
|||
Db 1 EVOLVESGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRTKNNVAT 60
QY 61 FYADSVKDRFTTSRDSSQSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYGQTSILT 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YYADSVKDRYTTISRDSKNTLYLQNNLKTEDTAMYYCVTFYNGV-----WGTGTVT 114
QY 121 VSS 123
|||
```

```
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 100
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-840-459-100

Query Match 80.3%; Score 519; DB 10; Length 148;
Best Local Similarity 80.5%; Pred. No. 6,9e-41;
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;

QY 1 EYWLVSGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRKSNVAT 60
|||
Db 20 EVOLVESGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRTKNNVAT 79
QY 61 FYADSVKDRFTTSRDSSQSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYGQTSILT 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 80 YYADSVKDRYTTISRDSSEMLFLQNNLKTEDTAMYYCVTFYNGV-----WGTGTVT 133
QY 121 VSS 123
|||
Db 134 VSS 136

RESULT 7
US-09-835-087-13
; Sequence 13, Application US/09835087
; Patent No. US20020042370A1
; GENERAL INFORMATION:
; APPLICANT: Wayne W. Hancock
; TITLE OF INVENTION: Method of Treating Graft Rejection Using
; FILE REFERENCE: 1855.2008-003
; CURRENT APPLICATION NUMBER: US/09/835,087
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/549,448
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: Humanized sequence
US-09-835-087-13

Query Match 76.3%; Score 493; DB 10; Length 117;
Best Local Similarity 76.4%; Pred. No. 1,3e-38;
Matches 94; Conservative 14; Mismatches 9; Indels 6; Gaps 1;

QY 1 EYWLVSGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRKSNVAT 60
|||
Db 1 EVOLVESGGGLVQPKGSLKLSCAASGFSPFNAYAMNVRQAPGKLEWVARIRTKNNVAT 60
QY 61 FYADSVKDRFTTSRDSSQSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYGQTSILT 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YYADSVKDRYTTISRDSKNTLYLQNNLKTEDTAMYYCVTFYNGV-----WGTGTVT 114
QY 121 VSS 123
|||
```

```

Db          115 VSS 117

RESULT 8
US-09-809-739-23
: Sequence 23, Application US/09809739
: Patent No. US20020106369A1
: GENERAL INFORMATION:
: APPLICANT: Horvath, Christopher J.
: APPLICANT: Rao, Patricia E.
: TITLE OF INVENTION: Method of Inhibiting Stenosis and
: TITLE OF INVENTION: Restenosis
: FILE REFERENCE: 1855.1069-003
: CURRENT APPLICATION NUMBER: US/09/809,739
: CURRENT FILING DATE: 2001-03-15
: PRIOR APPLICATION NUMBER: US 09/528,267
: PRIOR FILING DATE: 2000-03-17
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 23
: LENGTH: 117
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Humanized sequence
US-09-809-739-23

```

```
Query Match Similarity      76.3% Score 493 DB 10 Length 117  
Best Local Similarity      76.4% Pred NO 1.3e-38;  
Matches    94; Conservative   14; Mismatches     9; Indels       6; Gaps        1.  
  
OY           1 EVMVLVESSGGGLVQPKGSLKLSCAASGFTEFNMYAMMVROAPKGLEWVARIRSKSNNTAT 60  
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB          1 EVQLVESGGGLVKPCGSRLRSCAAGSFSPNAYMMWRQAPCKGLEWARIRFTKNNNMAT 60  
  
OY         61 FYADSVKRDRFTRSDSOSMLYLQNNNLKTETPTAMYCYCRARASGDIDAMDYGQTSLT 120  
            : | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB         61 FYYADVSRKYRTISRDSSKNITLYLNNSLKTEDTAIVYCCTFYNGNV-----WGQTLVT 114  
  
OY             121 VSS 123  
              |||  
DB            115 VSS 117  
  
RESULT 9  
US-09-840-459-20  
Sequence 20, Application US/09840459  
Patent No. US20020150576A1  
GENERAL INFORMATION:  
  APPLICANT: Larosa, Gregory J.  
  APPLICANT: Horvath, Christopher  
  APPLICANT: Newman, Walter  
  APPLICANT: Jones, S. Tarran  
  APPLICANT: O'Brien, Slobohan H.  
  TITLE OF INVENTION: HUMANIZED ANTI-CRCR ANTIBODIES AND  
  FILE REFERENCE: METHODS OF USE THEREFOR  
  CURRENT APPLICATION NUMBER: US/09/840.459  
  CURRENT FILING DATE: 2001-02-02  
  PRIOR APPLICATION NUMBER: PCT/US01/03537  
  PRIOR FILING DATE: 2001-02-02  
  PRIOR APPLICATION NUMBER: 09/497,625  
  PRIOR FILING DATE: 2000-02-03  
  PRIOR APPLICATION NUMBER: 09/359,193  
  PRIOR FILING DATE: 1999-07-22  
  PRIOR APPLICATION NUMBER: 09/121,781  
  PRIOR FILING DATE: 1998-07-23  
  NUMBER OF SEQ ID NOS: 107  
  SOFTWARE: FastSeq for Windows Version 3.0  
  LENGTH: 117  
  TYPE: PRT
```

```

; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Humanized sequence
US-09-840-459-20

```

Query Match	76.38;	Score 493;	DB 10;	length 117;
Best Local Similarity	-76.48;	Pred. No. 1.3e-38;		
Matches	94;	Conservative	14;	Mismatches 9;
			Indels	6;
			Gaps	1.

[illegible]

```

RESULT 10
US-09-835-087-12
Sequence 12, Application us/09835087
Patent No. US20020042370A1
GENERAL INFORMATION:
APPLICANT: Wayne W. Hancock
TITLE OF INVENTION: Method of Treating Graft Rejection Using
TITLE OF INVENTION: Inhibitors of CCR2 Function
FILE REFERENCE: 1855, 2008-003
CURRENT APPLICATION NUMBER: US/09/835,087
CURRENT FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: 09/549,448
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ. ID NOS.: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 117
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Humanized sequence
US-09-835-087-12

```

```

Query Match          75.7%; Score 489; DB 10; Length 117;
Best Local Similarity 75.6%; Pred. No. 3e-38;
Matches 93; Conservative 14; Mismatches 10; Indels 6; Gaps 1

OY      1 EYMLVESGGGLVQPAGSLKLSCAASGFPENNYAMNVRQAPGKGLWVARIRSKSNNTAT 60
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1 EQQLVESGGGLVKKPGGSLRLSCAASGFSFNAYAMNVRQAPGKGLEWVARIRTKNNNTAT 60
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      61 FYADSVKDRFTISRDDSSMLYLQNNLNKLTEDTAMYYCYVRCAASGIDVAMDYWGQSTSLT 120
      61 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      61 YYADSVKDRYTTISRDDSKNTLYLQNNLSKLTEDTAVYYCTTFYNGV-----WGQSTLVT 114
      61 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      121 VSS 123
      121 | | |
DB      115 VSS 117

RESULT 11
US-09-809-739-22
; Sequence 22, Application US/09809739
; Patent No. US20020106369A1
; GENERAL INFORMATION:
; APPLICANT: Horvath, Christopher J.
; APPLICANT: Rao, Patricia E.
; TITLE OF INVENTION: Method of Inhibiting Stenosis and
; TITLE OF INVENTION: Restenosis
; FILE REFERENCE: 1855.1069-003
; CURRENT APPLICATION NUMBER: US/09/809,739

```

QY 61 FYADSVKDRFTISRDDSQSMYLOMNNLKTEDTAMYCVRRGASGIDYAMDYWGQGSTLT 120

Query Match	75.5%; Score 488; DB 10; Length 117;
-------------	--------------------------------------

Best Local Similarity 75.6%; Pred. No. 3.7e-38;  
Matches 93; Conservative 13; Mismatches 11; Indels 6; Gaps 1;

```
QY 1 EVMIVESGGGLVOPKSGIKLSCAASGFTFNNYAMNWVROAPGKLEWVARIRSKSNYYAT 60
  |||
Db 1 EVOLVESGGGLVOPKSGIKLSCAASGFTFNNYAMNWVROAPGKLEWVARIRSKSNYYAT 60
  |||
QY 61 FYADSVADRFTISRDDSGSMILYLOMNNLKTEDTAMYYCVRGASGIDYAMDYWGQTSLT 120
  |||
Db 61 YYADSVADRFTISRDDSGSMILYLOMNNLKTEDTAVYYCTTFYNGV-----WGQGLVLT 114
  |||
QY 121 VSS 123
  |||
Db 115 VSS 117
```

```
RESULT 15
US-09-840-459-18
; Sequence 18, Application US/09840459
; Patent No. US20020150576a1
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Slobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 185.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanized sequence
US-09-840-459-18
```

```
Query Match 75.5%; Score 488; DB 10; Length 117;
Best Local Similarity 75.6%; Pred. No. 3.7e-38;
Matches 93; Conservative 13; Mismatches 11; Indels 6; Gaps 1;

QY 1 EVMIVESGGGLVOPKSGIKLSCAASGFTFNNYAMNWVROAPGKLEWVARIRSKSNYYAT 60
  |||
Db 1 EVOLVESGGGLVOPKSGIKLSCAASGFTFNNYAMNWVROAPGKLEWVARIRSKSNYYAT 60
  |||
QY 61 FYADSVADRFTISRDDSGSMILYLOMNNLKTEDTAMYYCVRGASGIDYAMDYWGQTSLT 120
  |||
Db 61 YYADSVADRFTISRDDSGSMILYLOMNNLKTEDTAVYYCTTFYNGV-----WGQGLVLT 114
  |||
QY 121 VSS 123
  |||
Db 115 VSS 117
```

Search completed: November 27, 2002, 07:26:00  
Job time : 15.2471 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:18:10 : Search time 1.73745 Seconds  
(Without alignments)  
137.479 Million cell updates/sec

Title: US-09-893-615-1  
Perfect score: 91  
Sequence: 1 WRMYFSHRHAHLRSP 15

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications: AA: \*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB pep: \*  
2: /cgn2\_6/ptodata/1/pubppaa/PCR\_NEW\_PUB pep: \*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB pep: \*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB pep: \*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB pep: \*  
6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB pep: \*  
7: /cgn2\_6/ptodata/1/pubppaa/PCRUS\_PUBCOMB pep: \*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB pep: \*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB pep: \*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB pep: \*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB pep: \*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB pep: \*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB pep: \*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	91	100.0	15	US-09-893-615-1	Sequence 1, Appl
2	91	100.0	19	US-09-893-615-27	Sequence 27, Appl
3	91	100.0	19	US-09-893-615-29	Sequence 29, Appl
4	91	100.0	19	US-09-893-615-33	Sequence 33, Appl
5	91	100.0	19	US-09-893-615-35	Sequence 35, Appl
6	91	100.0	19	US-09-893-615-37	Sequence 37, Appl
7	91	100.0	19	US-09-893-615-65	Sequence 65, Appl
8	66	72.5	19	US-09-893-615-31	Sequence 31, Appl
9	41	45.1	127	US-09-864-761-44155	Sequence 44155, A
10	41	45.1	191	US-09-828-644-87	Sequence 87, Appl
11	40	44.0	300	US-09-728-721-4	Sequence 4, Appl
12	40	44.0	300	US-10-105-931-4	Sequence 4, Appl
13	40	44.0	540	US-09-748-537-1	Sequence 1, Appl
14	40	44.0	540	US-09-728-721-2	Sequence 2, Appl
15	40	44.0	540	US-09-771-161A-184	Sequence 184, App
16	40	44.0	540	US-09-862-027-28	Sequence 28, Appl
17	40	44.0	540	US-10-133-780-1	Sequence 1, Appl
18	40	44.0	540	US-10-105-931-2	Sequence 2, Appl
19	40	44.0	544	US-09-925-301-1015	Sequence 1015, Ap

20	40	44.0	616	US-09-925-300-1519	Sequence 1519, Ap
21	39	42.9	92	US-10-032-159A-18	Sequence 18, Appl
22	39	42.9	139	US-10-032-159A-16	Sequence 16, Appl
23	39	42.9	378	US-09-864-761-43251	Sequence 43251, A
24	39	42.9	1004	US-09-767-215-2	Sequence 2, Appl
25	39	42.9	1138	US-09-767-215-5	Sequence 5, Appl
26	38	41.8	29	US-09-880-901-2	Sequence 2, Appl
27	38	41.8	39	US-09-864-761-48462	Sequence 48462, A
28	38	41.8	65	US-09-912-020-395	Sequence 395, App
29	38	41.8	65	US-09-915-242-10177	Sequence 10177, A
30	38	41.8	293	US-09-815-242-5614	Sequence 5614, A
31	38	41.8	303	US-09-828-523A-20	Sequence 20, Appl
32	38	41.8	303	US-09-815-242-12527	Sequence 12527, A
33	38	41.8	311	US-09-828-523A-86	Sequence 86, Appl
34	38	41.8	379	US-09-864-761-43222	Sequence 43222, A
35	37	40.7	29	US-09-995-297-34	Sequence 34, Appl
36	37	40.7	69	US-09-864-761-35570	Sequence 35570, A
37	37	40.7	69	US-09-864-761-49114	Sequence 49114, A
38	37	40.7	121	US-09-764-864-868	Sequence 868, App
39	37	40.7	189	US-09-987-967-2	Sequence 2, Appl
40	37	40.7	220	US-09-925-301-1209	Sequence 1209, Ap
41	37	40.7	224	US-09-885-188-11	Sequence 11, Appl
42	37	40.7	224	US-09-885-189-11	Sequence 11, Appl
43	37	40.7	267	US-09-726-643-63	Sequence 63, Appl
44	37	40.7	329	US-09-895-913A-170	Sequence 170, App
45	37	40.7	628	US-09-828-447-12	Sequence 12, Appl

#### ALIGNMENTS

RESULT 1  
US-09-893-615-1  
Sequence 1, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: ONSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Etinandi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995, 0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: Linear

MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 1  
US-09-893-615-1

Query Match 100.0%; Score 91; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.9e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRMYFSHRHAHLRSP 15  
Db 1 WRMYFSHRHAHLRSP 15

RESULT 2  
US-09-893-615-27  
Sequence 27, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:

APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Mong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315

## COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-893-615-27

Query Match 100.0%; Score 91; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.7e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRMYFSHRHAHLRSP 15  
Db 3 WRMYFSHRHAHLRSP 17

RESULT 3  
US-09-893-615-29  
Sequence 29, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.

Schuman, Richard F.  
Mong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315

## COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
US-09-893-615-29

Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 3.7e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WRMYFSHRHAHLRSP 15  
Db 3 WRMYFSHRHAHLRSP 17

RESULT 4  
US-09-893-615-33  
Sequence 33, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:

APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Mong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315

## COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/893.615
:   FILING DATE: 29-Jun-2001
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Elnaudt, Carol P.
:     REGISTRATION NUMBER: 32,220
:     REFERENCE/DOCKET NUMBER: 04995.0041-00000
:   TELECOMMUNICATION INFORMATION:
:     TELEPHONE: 202-408-4400
:     TELEFAX: 202-408-4400
:   INFORMATION FOR SEQ ID NO: 33:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 19 amino acids
:       TYPE: amino acid
:       TOPOLOGY: linear
:   MOLECULE TYPE: protein
:   SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-893-615-33

Query Match      100.0%; Score 91; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 WRWFSHRHAHLRSP 15
        |||
Db      3 WRWFSHRHAHLRSP 17

RESULT 5
US-09-893-615-35
: Sequence 35, Application US/09893615
: Patent No. US20020082395A1
: GENERAL INFORMATION:
:   APPLICANT: Fischer, Gerald W.
:             Schuman, Richard F.
:             Wong, Hing
:             Stinson, Jeffrey L.
:   TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND
:                     CHIMERIC ANTIBODIES SPECIFIC FOR LIPOPEPTIC ACID OF GRAM
:                     POSITIVE BACTERIA
:   NUMBER OF SEQUENCES: 89
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
:               STREET: 1300 I Street, NW
:               CITY: Washington
:               STATE: DC
:               COUNTRY: USA
:               ZIP: 20005-3315
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: PatentIn Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/09/893.615
:     FILING DATE: 29-Jun-2001
:     ATTORNEY/AGENT INFORMATION:
:       NAME: Elnaudt, Carol P.
:       REGISTRATION NUMBER: 32,220
:       REFERENCE/DOCKET NUMBER: 04995.0041-00000
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: 202-408-4400
:       TELEFAX: 202-408-4400
:     INFORMATION FOR SEQ ID NO: 35:
:       SEQUENCE CHARACTERISTICS:
:         LENGTH: 19 amino acids
:         TYPE: amino acid
:         TOPOLOGY: linear
:       MOLECULE TYPE: protein
:       SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-893-615-35

Query Match      100.0%; Score 91; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 WRWFSHRHAHLRSP 15
        |||
Db      3 WRWFSHRHAHLRSP 17

RESULT 6
US-09-893-615-37
: Sequence 37, Application US/09893615
: Patent No. US20020082395A1
: GENERAL INFORMATION:
:   APPLICANT: Fischer, Gerald W.
:             Schuman, Richard F.
:             Wong, Hing
:             Stinson, Jeffrey L.
:   TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND
:                     CHIMERIC ANTIBODIES SPECIFIC FOR LIPOPEPTIC ACID OF GR
:                     POSITIVE BACTERIA
:   NUMBER OF SEQUENCES: 89
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
:               STREET: 1300 I Street, NW
:               CITY: Washington
:               STATE: DC
:               COUNTRY: USA
:               ZIP: 20005-3315
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: PatentIn Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/09/893.615
:     FILING DATE: 29-Jun-2001
:     ATTORNEY/AGENT INFORMATION:
:       NAME: Elnaudt, Carol P.
:       REGISTRATION NUMBER: 32,220
:       REFERENCE/DOCKET NUMBER: 04995.0041-00000
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: 202-408-4400
:       TELEFAX: 202-408-4400
:     INFORMATION FOR SEQ ID NO: 37:
:       SEQUENCE CHARACTERISTICS:
:         LENGTH: 19 amino acids
:         TYPE: amino acid
:         TOPOLOGY: linear
:       MOLECULE TYPE: protein
:       SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-893-615-37

Query Match      100.0%; Score 91; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 WRWFSHRHAHLRSP 15
        |||
Db      3 WRWFSHRHAHLRSP 17

RESULT 7
US-09-893-615-65
: Sequence 65, Application US/09893615
: Patent No. US20020082395A1
: GENERAL INFORMATION:
:   APPLICANT: Fischer, Gerald W.
:             Schuman, Richard F.
:             Wong, Hing
:             Stinson, Jeffrey L.
:   TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND
:                     CHIMERIC ANTIBODIES SPECIFIC FOR LIPOPEPTIC ACID OF GR
:                     POSITIVE BACTERIA
:   NUMBER OF SEQUENCES: 89
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
:               STREET: 1300 I Street, NW
:               CITY: Washington
:               STATE: DC
:               COUNTRY: USA
:               ZIP: 20005-3315
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: PatentIn Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/09/893.615
:     FILING DATE: 29-Jun-2001
:     ATTORNEY/AGENT INFORMATION:
:       NAME: Elnaudt, Carol P.
:       REGISTRATION NUMBER: 32,220
:       REFERENCE/DOCKET NUMBER: 04995.0041-00000
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: 202-408-4400
:       TELEFAX: 202-408-4400
:     INFORMATION FOR SEQ ID NO: 65:
:       SEQUENCE CHARACTERISTICS:
:         LENGTH: 19 amino acids
:         TYPE: amino acid
:         TOPOLOGY: linear
:       MOLECULE TYPE: protein
:       SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-09-893-615-65
```

TITLE OF INVENTION: OPOSonic AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: Version #1.0, Version #1.30  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4400  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-09-893-615-65  
Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 3.7e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 3 WRMYFSHRHAHLRSP 17  
1 WRMYFSHRHAHLRSP 15  
|||||  
US-09-893-615-31  
RESULT 8  
Sequence 31, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPOSonic AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: Version #1.0, Version #1.30  
APPLICATION NUMBER: US/09/893,615

FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4400  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-09-893-615-31  
Query Match 72.5%; Score 66; DB 10; Length 19;  
Best Local Similarity 73.3%; Pred. No. 0.00023;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
DB 3 WRMYFSHRHAHLRSP 17  
1 WRMYFSHRHAHLRSP 15  
|||||  
DB 3 WRMYFSHRHAHLRSP 17  
RESULT 9  
US-09-864-761-44155  
Sequence 44155, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmlca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408

```

1  PRIOR FILING DATE: 2000-06-30
2  PRIOR APPLICATION NUMBER: 09/774,203
3  PRIOR FILING DATE: 2001-01-29
4  NUMBER OF SEQ. ID NOS: 49117
5  SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
6  SEQ. ID NO 44135
7  LENGTH: 127
8  TYPE: PRT
9  ORGANISM: Homo sapiens
10 FEATURE:
11 OTHER INFORMATION: MAP TO AC004622.1
12 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.72
13 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
14 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.71
15 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.79
16 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
17 OTHER INFORMATION: EST_HUMAN HIT: AW0502362.1, EVALU0 5.00e-40
18 US-09-864-761-44135

```

```
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 540
; TYPE: prt
; ORGANISM: Homo sapiens
; US-09-748-537-1
```

```
Query Match          44.0%; Score 40; DB 10; Length 540;
Best Local Similarity 33.3%; Pred. No. 57;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 WRMYFSHRHAHLRSP 15
   ||: : ||: : ||
Db 40 WRQYAVAKHLHTP 54
```

```
RESULT 14
US-09-728-721-2
; Sequence 2, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 540
; TYPE: prt
; ORGANISM: Homo sapiens
; US-09-728-721-2
```

```
Query Match          44.0%; Score 40; DB 10; Length 540;
Best Local Similarity 33.3%; Pred. No. 57;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 WRMYFSHRHAHLRSP 15
   ||: : ||: : ||
Db 40 WRQYAVAKHLHTP 54
```

```
RESULT 15
US-09-771-161A-184
; Sequence 184, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
```

```
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 184
; LENGTH: 540
; TYPE: prt
; ORGANISM: Homo sapiens
; US-09-771-161A-184
```

```
Query Match          44.0%; Score 40; DB 10; Length 540;
Best Local Similarity 33.3%; Pred. No. 57;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 WRMYFSHRHAHLRSP 15
   ||: : ||: : ||
Db 40 WRQYAVAKHLHTP 54
```

```
Search completed: November 27, 2002, 07:25:58
Job time : 1.73745 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:18:10 : Search time 1.73745 Seconds  
(without alignments)  
137.479 Million cell updates/sec

Title: US-09-893-615-2

Perfect score: 91

Sequence: 1 WHMRRIRPLQLAGR 15

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	15	US-09-893-615-2	Sequence 2, Appl
2	91	100.0	19	US-09-893-615-9	Sequence 9, Appl
3	91	100.0	19	US-09-893-615-17	Sequence 17, Appl
4	91	100.0	19	US-09-893-615-19	Sequence 19, Appl
5	91	100.0	19	US-09-893-615-41	Sequence 41, Appl
6	91	100.0	19	US-09-893-615-57	Sequence 57, Appl
7	80	87.9	19	US-09-893-615-15	Sequence 15, Appl
8	50	54.9	139	US-09-864-761-33721	Sequence 33721, A
9	44.5	48.9	323	US-09-816-087-4	Sequence 4, Appl
10	44.5	48.9	343	US-09-985-694A-9	Sequence 9, Appl
11	43	47.3	51	US-09-864-761-42937	Sequence 42937, A
12	43	47.3	1298	US-09-982-610-33	Sequence 33, Appl
13	42	46.2	68	US-09-920-920-81	Sequence 81, Appl
14	41	45.1	54	US-09-925-297-466	Sequence 466, App
15	41	45.1	99	US-09-864-761-40505	Sequence 40505, A
16	41	45.1	379	US-09-804-682-72	Sequence 72, Appl
17	41	45.1	379	US-09-804-682-73	Sequence 73, Appl
18	41	45.1	411	US-09-804-682-74	Sequence 74, Appl
19	40	44.0	54	US-09-925-297-888	Sequence 888, App

20	40	44.0	136	US-09-728-479-9	Sequence 9, Appl
21	40	44.0	136	US-09-263-689-13	Sequence 13, Appl
22	40	44.0	802	US-09-964-899-41	Sequence 41, Appl
23	40	44.0	1074	US-09-509-106A-2	Sequence 2, Appl
24	39.5	43.4	557	US-09-798-743A-1	Sequence 1, Appl
25	39	42.9	96	US-09-864-761-45697	Sequence 45697, A
26	39	42.9	253	US-09-815-242-11842	Sequence 11842, A
27	39	42.9	338	US-09-318-271-8	Sequence 8, Appl
28	38.5	42.3	343	US-09-985-694A-2	Sequence 2, Appl
29	38	41.8	25	US-09-864-761-43467	Sequence 43467, A
30	38	41.8	58	US-09-864-761-48958	Sequence 48958, A
31	38	41.8	133	US-09-867-550-372	Sequence 372, App
32	38	41.8	577	US-09-815-242-11760	Sequence 11760, A
33	38	41.8	797	US-09-815-242-5071	Sequence 5071, Ap
34	38	41.8	824	US-09-932-145-3	Sequence 3, Appl
35	37.5	41.2	189	US-09-815-242-11799	Sequence 11799, A
36	37.5	41.2	264	US-09-815-242-13980	Sequence 13980, A
37	37.5	41.2	417	US-09-815-242-10165	Sequence 10165, A
38	37	40.7	73	US-09-764-887-288	Sequence 288, App
39	37	40.7	79	US-09-867-550-860	Sequence 860, App
40	37	40.7	178	US-09-966-546-2	Sequence 2, Appl
41	37	40.7	178	US-09-966-545-2	Sequence 2, Appl
42	37	40.7	610	US-09-748-107-2	Sequence 2, Appl
43	36.5	40.1	295	US-09-815-242-10104	Sequence 10104, A
44	36.5	40.1	557	US-09-798-743A-3	Sequence 3, Appl
45	36	39.6	11	US-09-780-070-21	Sequence 21, Appl

#### ALIGNMENTS

RESULT 1  
US-09-893-615-2  
Sequence 2, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischeer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GF  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESS: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: Linear

MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 2  
US-09-893-615-2

Query Match 100.0%; Score 91; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 7.7e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPLQLAGR 15  
Db 1 WHMRHRIPLQLAGR 15

## RESULT 2

US-09-893-615-9  
Sequence 9, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA

NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001

ATTORNEY/AGENT INFORMATION:  
NAME: Einaudt, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-09-893-615-9

Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPLQLAGR 15  
Db 3 WHMRHRIPLQLAGR 17

## RESULT 3

US-09-893-615-17  
Sequence 17, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.

Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA

NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001

ATTORNEY/AGENT INFORMATION:  
NAME: Einaudt, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-893-615-17

Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMRHRIPLQLAGR 15  
Db 3 WHMRHRIPLQLAGR 17

## RESULT 4

US-09-893-615-19  
Sequence 19, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.

TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA

NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001



SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-893-615-19

Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. NO. 9.6e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHRRHRIPQLAAGR 15  
Db 3 WHRRHRIPQLAAGR 17

RESULT 5  
US-09-893-615-41  
Sequence 41, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Mong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 41:  
US-09-893-615-41

Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. NO. 9.6e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHRRHRIPQLAAGR 15  
Db 3 WHRRHRIPQLAAGR 17

RESULT 6  
US-09-893-615-57  
Sequence 57, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Mong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 57:  
US-09-893-615-57

Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. NO. 9.6e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WHRRHRIPQLAAGR 15  
Db 3 WHRRHRIPQLAAGR 17

RESULT 7  
US-09-893-615-15  
Sequence 15, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Mong, Hing  
Stinson, Jeffrey L.

TITLE OF INVENTION: OPOSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elnaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-893-615-15  
Query Match 87.9%; Score 80; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 4e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 HMRHRIPLOLAGR 15  
DB 4 HMRHRIPLOLAGR 17  
RESULT 8  
US-09-864-761-33721  
Sequence 33721, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 33721  
LENGTH: 139  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AP000046.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.2  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.3  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.3  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.8  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5  
OTHER INFORMATION: EST\_HUMAN HIT: AA853678.1, EVALUATE 2.00e-21  
US-09-864-761-33721  
Query Match 54.9%; Score 50; DB 10; Length 139;  
Best Local Similarity 72.7%; Pred. No. 0.68;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 WHMRHRIPLOL 11  
DB 98 WHMRHRIPLOL 108  
RESULT 9  
US-09-816-087-4  
Sequence 4, Application US/09816087  
Patent No. US20020064822A1  
GENERAL INFORMATION:  
APPLICANT: Wei, Ming-Hui  
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
FILE REFERENCE: CL000749-CIP  
CURRENT APPLICATION NUMBER: US/09/816,087  
CURRENT FILING DATE: 2001-03-26  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 323  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-816-087-4  
Query Match 48.9%; Score 44.5; DB 10; Length 323;

Best Local Similarity	57.18;	Pred. No.	9.7;
Matches	8;	Conservative	3;
		Mismatches	2;
		Indels	1;
		Gaps	1;

```

Oy      1 WHWRHRIPQLAAG 14
          | : | | | : | : | |
Db      138 WYWRRR-PKRLSAG 150

```

RESULT 10  
US-09-985-694A-9

```

? TITLE OF INVENTION: G-Protein Coupled Receptor
? FILE REFERENCE: PFI45P1
? CURRENT APPLICATION NUMBER: US/09/985,694A
? PRIOR FILING DATE: 2001-11-05
? PRIOR APPLICATION NUMBER: 08/461,989
? PRIOR FILING DATE: 1995-06-05
? PRIOR APPLICATION NUMBER: PCT/US94/13296
? PRIOR FILING DATE: 1994-11-18
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: PatentIn version 3.1

```

Query Match	48.9%	Score 44.5	DB 10	Length 343
Best Local Similarity	57.1%	Pred. NO. 10		
Matches	8	Conservative	3	Mismatches 2
				Indels 1
				Gaps 1

```

QY      1 WHWRHRIPQLAAG 14
          |::| | :|::|
Db      150 WYWRRR-PKRLSAG 162

```

RESULT 11  
US-09-864-761-42937  
; Sequence 42937, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:

: APPLICANT: Penn, Sharon G.  
 : APPLICANT: Rank, David R.  
 : APPLICANT: Hanzel, David K.  
 : APPLICANT: Chen, Wensheng  
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 : TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

PRIORITY APPLICATION NUMBER: PCT/US01/006668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/006663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/006662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/006661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/006670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomach Sequence Listing Engine vers. 1.1
SEQ ID NO 42937
LENGTH: 51
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL049594.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.9
OTHER INFORMATION: EST_HUMAN HIT: BF338634.1, EVALUATE 4.00e-2
OTHER INFORMATION: SWISSPROT HIT: P42128, EVALUATE 5.20e-01
US-09-864 -761-42937 .

```

Query Match	47.3%	Score 43	DB 10	Length 51
Best Local Similarity	44.4%	Pred. No. 2.8		
Matches	8	Conservative	1	Indels 8
				Gaps 1

```
QY      1 WHWR-----HRIPLQ 10
          | | | | |
Db      30 WDWRLCGRGGRPHRLPLQ 47
```

RESULT 12  
US-09-982-610-33  
; Sequence 33, Application US/09982610  
; Patent No. US20020146420A1

APPLICANT: Genentech, Inc.  
Bennett, Brian D.  
Goeddel, David  
Lee, James M.  
Mathews, William  
Tsai, Siao Ping  
Wood, William I.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,610  
FILING DATE: 17-Oct-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,648  
FILING DATE: 1996-MAY-23  
APPLICATION NUMBER: 08/222616

```

: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M.
: REGISTRATION NUMBER: 40,378
: REFERENCE/DOCKET NUMBER: P0821P3PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 33:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1298 amino acids
: TYPE: Amino Acid
: TOPOLOGY: Linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-982-610-33

Query Match 47.3%; Score 43; DB 10; Length 1298;
Best Local Similarity 46.7%; Pred. No. 61;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0.

OY 1 WHHRRIPQLAGR 15
||||| |::| |
Db 458 WHHRRIPCKMPAQR 472

RESULT 13
US-09-220-920-81
: Sequence 81, Application US/09220920
: Patent No. US2002002269A1
: GENERAL INFORMATION:
: APPLICANT: Mlbrandt, Jeffrey D.
: APPLICANT: Baloh, Robert H.
: TITLE OF INVENTION: Artemin, A No. US2002002269A1el Neurotrophic Factor
: FILE REFERENCE: 6029-7996
: CURRENT APPLICATION NUMBER: US/09/220,920
: CURRENT FILING DATE: 1998-12-24
: EARLIER APPLICATION NUMBER: 09/163,283
: EARLIER FILING DATE: 1998-09-29
: EARLIER APPLICATION NUMBER: 60/108,148
: EARLIER FILING DATE: 1998-11-12
: EARLIER APPLICATION NUMBER: 09/218,698
: EARLIER FILING DATE: 1998-12-22
: NUMBER OF SEQ ID NOS: 120
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 81
: LENGTH: 68
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-220-920-81

Query Match 46.2%; Score 42; DB 10; Length 68;
Best Local Similarity 57.1%; Pred. No. 5.1;
Matches 8; Conservative 1; Mismatches 3; Indels 2; Gaps 1.

OY 1 WHHRR--HRIPLGA 12
||||| |::| |
Db 55 WTRPLHAYPLPLA 68

RESULT 14
US-09-925-297-466
: Sequence 466, Application US/09925297
: Patent No. US20020081659A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA105
: CURRENT APPLICATION NUMBER: US/09/925,297
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/0500/05969
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270

```

```

: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 928
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 466
: LENGTH: 54
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-925-297-466

Query Match      45.18;   Score 41;   DB 10;   Length 54;
Best Local Similarity 58.3%;   Pred. No. 5.8;
Matches 7;   Conservative 1;   Mismatches 4;   Indels 0;   Gaps 0;

Oy      2   HWRHRIPLOLAA 13
          |||||
Db      9   HWRHRNPILKATA 20

RESULT 15
US-09-864-761-40505
: Sequence 40505, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 40505
: LENGTH: 99
: TYPE: PRT

```

; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC013627.3.  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5  
; OTHER INFORMATION: SWISSPROT HIT: P28476, EVALU2.50e+00  
; OTHER INFORMATION: EST\_HUMAN HIT: AA843600.1, EVALU2.60e-01  
US-09-864-761-40505

Query Match 45.18; Score 41; DB 10; Length 99;  
Best Local Similarity 43.88; Pred. No. 10;  
Matches 7; Conservative 3; Mismatches 4; Indels 2; Gaps 1;  
OY 1 WHRHR--IPLQLAG 14  
|||:| ||:|  
DB 57 WHMNRKFAPLRMLPG 72

Search completed: November 27, 2002, 07:25:59  
Job time : 2.73745 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:23:05 : Search time 1.44788 Seconds

(without alignments)  
429.695 Million cell updates/sec

Title: US-09-893-615-2

Sequence: 1 MWMHRHRLQLAGLR 15

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	40.0	203	FLA1_ARCFU	O29208 archaeoglob
2	6	40.0	342	YM22_YEAST	P40214 saccharomyc
3	6	40.0	367	DP3B_PSEPU	P13455 pseudomonas
4	6	40.0	510	GARD_BACSU	P42240 bacillus su
5	6	40.0	536	CBP8_HUMAN	P22792 homo sapien
6	6	40.0	558	BCHD_RHOSH	O34845 rhodobacter
7	6	40.0	666	Y079_SCHPO	O13670 schizosacch
8	6	40.0	778	FTSK_COXBU	P39920 coxiella bu
9	6	40.0	1144	DP3A_NEIMA	O91928 neisseria m
10	6	40.0	1144	DP3A_NEIMA	O91928 neisseria m
11	5	33.3	229	HOXY_RHOOP	P22660 rhodococcus
12	5	33.3	93	REP1_ECOLI	P05830 escherichia
13	5	33.3	98	FER1_CVAPA	P17007 cyanophora
14	5	33.3	99	FER2_APHSA	P00251 aphanothece
15	5	33.3	102	YAIY_ECOLI	P77669 escherichia
16	5	33.3	105	YB59_YEAST	P38311 saccharomyc
17	5	33.3	108	NIFH_NOSMU	O09158 nostoc musc
18	5	33.3	112	OL11_MOUSE	O60890 mus musculu
19	5	33.3	114	Y118_ARCFU	O30118 archaeoglob
20	5	33.3	136	YAGP_ECOLI	P75684 escherichia
21	5	33.3	148	MBB1_ECOLI	P08097 escherichia
22	5	33.3	156	COAD_HAETN	P44805 haemophilus
23	5	33.3	157	VE6_HPV36	P50810 human papil
24	5	33.3	157	YB5_YEAST	O05027 saccharomyc
25	5	33.3	158	TPC1_BALNU	P21797 balanus nub
26	5	33.3	162	UL33_HSVB	P28953 equine herp
27	5	33.3	170	YFCR_ECOLI	P76501 escherichia
28	5	33.3	171	MSG2_MOUSE	P47759 mus musculu
29	5	33.3	179	HSLV_LISIN	O32C74 listeria in
30	5	33.3	179	HSLV_LISMO	O8Y719 listeria mo
31	5	33.3	179	YAJI_ECOLI	P61122 escherichia
32	5	33.3	180	HSLV_BACDH	O8K426 bacillus ha
33	5	33.3	181	HSLV_BACSU	P39070 bacillus su

34	5	33.3	184	YDJA_HAETN	P45244 haemophilus
35	5	33.3	192	MOBA_NEIMB	P58747 neisseria m
36	5	33.3	195	AANT_HDYM1	P25881 hepatitis d
37	5	33.3	196	UNPA_SALTY	P27667 salmonella
38	5	33.3	201	PCP_BACDH	O9K604 bacillus ha
39	5	33.3	203	IAAS_HORVU	P07596 hordeum vul
40	5	33.3	204	KGUA_XYLEFA	O9pd76 xylella fas
41	5	33.3	205	YICG_ECOLI	P31432 escherichia
42	5	33.3	210	Y593_CAME	O9phs3 campylobact
43	5	33.3	213	ISPD_AQUAE	O67343 aquilex aeo
44	5	33.3	215	HEX8_ADEMI	P19722 mouse adeo
45	5	33.3	215	TRPF_BACSU	P20167 bacillus su

## ALIGNMENTS

```

RESULT 1
FLA1_ARCFU STANDARD: PRT: 203 AA.
AC O29208:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable flagellin 1 precursor.
GN FLA1 OR AF1054.
OS Archaeoglobus fulgidus.
OC Archaeae; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=96049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kiknes E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujil C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF FLAGELLA.
CC -!- SIMILARITY: BELONGS TO THE ARCHAEAL FLAGELLIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001030; AAB90186.1; -
CC TIGR: AF1054;
CC InterPro: IPR002774; Arch_flagellin.
CC Pfam: PF01917; Arch_flagellin; 1.
CC KEGG: K01917; Arch_flagellin; 1.
CC PROTEIN: Complete proteome.
CC CHAIN: 1 203 POTENTIAL.
CC SEQUENCE 203 AA: 21865 MM; C1C9B1210D79809D C1C64;
Query Match 40.0%; Score 6; DB 1; Length 203;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 9 LQLAAG 14  
 |||||  
 Db 85 LQLAAG 90

## RESULT 2

YM22\_YEAST STANDARD: PRT: 342 AA.

AC P40214;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Hypothetical 38.6 kDa protein in RPS16A-TIF34 intergenic region.  
 GN YMR144W OR YMR375.13.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;

RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288c / AB972;  
 RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 247071; CAA87358.1; -  
 DR SGD: S0004752; YMR144W.  
 KW Hypothetical protein.  
 SQ SEQUENCE 342 AA: 38570 MW: AEA7FESEI080CA6 CRC64:

Query Match 40.0%; Score 6; DB 1; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAAG 14  
 |||||  
 Db 274 LQLAAG 279

## RESULT 3

DP3B\_PSEPU STANDARD: PRT: 367 AA.

AC P13455;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE DNA polymerase III, beta chain (EC 2.7.7.7).  
 GN DNAN.

OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 OX NCBI\_TaxID=303;

RN (1)  
 RP SEQUENCE FROM N.A.

RC STRAIN=TN2100;  
 RX MEDLINE=89218947; PubMed=2540413;

RA Fujita M.O., Yoshikawa H., Ogasawara N.;  
 RT "Structure of the dna region of Pseudomonas putida: conservation  
 RT among three bacteria, Bacillus subtilis, Escherichia coli and P.

RT putida.";  
 RL Mol. Genet. 215:381-387(1989).

CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME  
 CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.  
 CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.  
 CC THE BETA CHAIN IS REQUIRED FOR INITIATION OF REPLICATION ONCE IT  
 CC IS CLAMPED ONTO DNA, IT SLIDES FREELY (BIDIRECTIONAL AND ATP-

CC INDEPENDENT) ALONG DUPLEX DNA (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate;  
 CC + [DNA](N).  
 CC -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND  
 CC THETA CHAINS) THAT CAN REPAIR SHORT GAPS CREATED BY NUCLEASE  
 CC IN DUPLEX DNA. FOR EFFICIENT REPLICATION OF THE LONG, SINGLE-  
 CC STRANDED TEMPLATES, POL III REQUIRES THE AUXILIARY CHAINS BETA,  
 CC GAMMA, AND DELTA.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X14791; CAA32894.1; -  
 DR PIR: JY0002; DIPS3P.

DR HSSP: P00583; 2POL.  
 DR InterPro: IPR001001; DNA\_polIII\_beta.

DR Pfam: PF02767; DNA\_pol3\_beta.1.

DR Pfam: PF02768; DNA\_pol3\_beta.3; 1.

DR SMART: SM00480; POL3BC; 1.

DR TRIGRAMS: TRGR00663; dnan; 1.

DR Transference: DNA-directed DNA polymerase; DNA replication.  
 SQ SEQUENCE 367 AA: 40718 MW: FE65ABABAD65FEAD CRC64:

Query Match 40.0%; Score 6; DB 1; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAAG 14  
 |||||  
 Db 284 LQLAAG 289

## RESULT 4

GARD\_BACSU STANDARD: PRT: 510 AA.

AC P42240; P42241.  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable D-galactarate dehydratase (EC 4.2.1.42) (calco).

GN GARD.

OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

RN (1)  
 RP SEQUENCE FROM N.A.

RC STRAIN=168;  
 RX MEDLINE=95219079; PubMed=7704254;

RA Ogawa K.-I., Akagawa E., Nakamura K., Yamane K.;  
 RT "Determination of a 21548 bp nucleotide sequence around the 24  
 RT degrees region of the Bacillus subtilis chromosome.";  
 RL Microbiology 141:269-275(1995).

RN (12)  
 RP SEQUENCE FROM N.A.

RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alliont G.,  
 RA Azevedo V., Bertoro M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borrius R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruchl C.V., Caldwell B., Capano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,

RA Dettl K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Funo S., Galizzi A., Galleron N.,

RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,



RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasanara Y., Klaerli-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetle D., Rapoport G., Rey M., Reynolds S.,  
 RA Pirescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowski A., Seror S.J., Serro P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarelli A.,  
 RA Viari A., Mamput R., Wedler E., Wedler E., Wetzsteiger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT subtilis".  
 RT Nature 390:249-256(1997).  
 CC -1- CATALYTIC ACTIVITY: D-galactarate = 5-dehydro-4-deoxy-D-glucarate  
 CC + H2O.  
 CC -1- PATHWAY: FIRST STEP IN THE METABOLISM OF D-GALACTARATE.  
 CC -1- SIMILARITY: BELONGS TO THE UXA FAMILY.  
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO  
 CC FRAMESHIFTS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D30808; BAA06472.1; ALT\_FRAME.  
 DR EMBL: D30808; BAA06473.1; ALT\_FRAME.  
 DR EMBL: Z99105; CAB12045.1; -.  
 DR Subtilisin: BG1163; gard.  
 KW Lyase; Complete proteome.  
 SQ SEQUENCE 510 AA; 54792 MW; C7283A007174CDE3 CRC64;

Query Match 40.0%; Score 6; DB 1; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LQLAG 14  
 111111  
 DB 412 LQLAG 417

RESULT 5  
 CBB8\_HUMAN STANDARD; PRT; 536 AA.  
 ID P22792;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Carboxypeptidase N 83 kDa chain (Carboxypeptidase N regulatory  
 DE subunit) (Fragment).  
 GN CPN2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=90094386; PubMed=2378615;  
 RA Tan F., Weerasinghe D.K., Skidgel R.A., Tamei H., Kaul R.K.,  
 RA Roninson I.B., Schilling J.W., Erdoes E.G.;  
 RT "The deduced protein sequence of the human carboxypeptidase N high

RT molecular weight subunit reveals the presence of leucine-rich tandem  
 RT repeats.";  
 RL J. Biol. Chem. 265:13-19(1990).  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=88309120; PubMed=3408501;  
 RA Skidgel R.A., Bennett C.D., Schilling J.W., Tan F., Weerasinghe D.K.,  
 RA Erdoes E.G.;  
 RT "Amino acid sequence of the N-terminus and selected tryptic peptides  
 RT of the active subunit of human plasma carboxypeptidase N: comparison  
 RT with other carboxypeptidases".  
 RL Biochem. Biophys. Res. Commun. 154:1323-1329(1988).  
 CC -1- FUNCTION: THE 83 kDa SUBUNIT BINDS AND STABILIZES THE CATALYTIC  
 CC SUBUNIT AT 37 DEGREES CELSIUS AND KEEPS IT IN CIRCULATION. UNDER  
 CC SOME CIRCUMSTANCES IT MAY BE AN ALLOSTERIC MODIFIER OF THE  
 CC CATALYTIC SUBUNIT.  
 CC -1- SUBUNIT: TETRAMER OF TWO CATALYTIC CHAINS AND TWO GLYCOSYLATED  
 CC INACTIVE CHAINS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: O-GLYCOSYLATED IN THE SER/THR-RICH REGION (POTENTIAL).  
 CC -1- PTM: WHETHER OR NOT ANY CYS RESIDUES PARTICIPATE IN INTRACHAIN  
 CC BONDS IS UNKNOWN, BUT THEY DO NOT FORM INTERCHAIN DISULFIDE BONDS  
 CC WITH THE 50 kDa CATALYTIC SUBUNIT.  
 CC -1- DISEASE: A COMPLETE ABSENCE OF THE ENZYME IS NOT CONSIDERED TO BE  
 CC COMPATIBLE WITH LIFE.  
 CC -1- SIMILARITY: CONTAINS 13 LEUCINE-RICH REPEATS (LRR).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J05158; AAA51921.1; -.  
 DR PIR: A34901; A34901.  
 DR Genew: HGNC:2313; CPN2.  
 DR MIM: 603104; -.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000372; LRR\_Nterm.  
 DR InterPro: IPR003591; LRR\_Lyp.  
 DR Pfam: PF00560; LRR\_11.  
 DR SMART: SM00013; LRRNT; 1.  
 DR SMART: SM00369; LRR\_TYP; 10.  
 KW Repeat; Leucine-rich repeat; Glycoprotein.  
 KM NON\_TER  
 FT 1  
 FT REPEAT 51 74 LRR 1.  
 FT REPEAT 75 98 LRR 2.  
 FT REPEAT 99 122 LRR 3.  
 FT REPEAT 123 146 LRR 4.  
 FT REPEAT 148 170 LRR 5.  
 FT REPEAT 171 194 LRR 6.  
 FT REPEAT 196 218 LRR 7.  
 FT REPEAT 220 242 LRR 8.  
 FT REPEAT 243 266 LRR 9.  
 FT REPEAT 268 290 LRR 10.  
 FT REPEAT 291 314 LRR 11.  
 FT REPEAT 316 338 LRR 12.  
 FT REPEAT 339 362 LRR 13.  
 FT REPEAT 359 379 SER/THR-RICH.  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 536 AA; 58649 MW; D524011243C9B2A6 CkC64;

Query Match 40.0%; Score 6; DB 1; Length 536;  
 Best Local Similarity 100.0%; Pred. No. 32;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 8 PLOLAA 13  
Db 384 PLOLAA 389

## RESULT 6

BCHD\_RHOSH STANDARD: PRT: 558 AA.  
ID BCHE\_RHOSH  
AC O34845;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Magnesium-chelatase 60 kDa subunit (Mg-protoporphyrin IX chelatase)  
DE (Mg-chelatase subunit D).  
GN BCHE.  
OS Rhodospirillum rubrum (Rhodospirillum rubrum).  
OC Bacteria: Proteobacteria; alpha subdivision; Rhodospirillum group.  
OC Rhodospirillum.  
OX NCBI\_TaxID=1063;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;  
RA Hansson M., Kannangara C.G.;  
RT "Rhodospirillum rubrum bchl and bchd encoding two subunits of  
RT magnesium chelatase.";  
RL (in) Plant Gene Register PGR97-190.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;  
RX MEDLINE=20115911; PubMed=10648776;  
RA Choudhary M., Kaplan S.;  
RT "DNA sequence analysis of the photosynthesis region of Rhodospirillum  
RT rubrum 2.4.1.";  
RL Nucleic Acids Res. 28:862-867 (2000).  
RN [3]  
RP CHARACTERIZATION.  
RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;  
RX MEDLINE=95199269; PubMed=7892204;  
RA Gibson L.C.D., Willows R.D., Kannangara C.G., von Wettstein D.,  
RA Hunter C.N.;  
RT "Magnesium-protoporphyrin chelatase of Rhodospirillum rubrum:  
RT reconstitution of activity by combining the products of the bchl, -I,  
RT and -D genes expressed in Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:11941-11944 (1995).  
CC -1- FUNCTION: Involved in bacteriochlorophyll biosynthesis; introduces  
CC a magnesium ion into protoporphyrin IX to yield Mg-protoporphyrin  
CC IX.  
CC -1- PATHWAY: Bacteriochlorophyll biosynthesis.  
CC -1- SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 VMFA DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC DR EMBL: AJ010302; CAB38736.1;  
CC DR EMBL: AF017642; AB97157.1;  
CC DR EMBL: AF195122; AF24286.1;  
CC DR InterPro: IPR002035; VMFA\_A.  
CC DR SMART: SM00327; VMFA\_1.  
CC DR PROSITE: PS50234; VMFA\_1.  
CC KM Photosynthesis: Bacteriochlorophyll biosynthesis.  
CC FT DOMAIN 376 VMFA.  
CC FT DOMAIN 238 272 GIU/PRO-RICH.  
CC SO SEQUENCE 558 AA; 58548 MW; 2A16F62C1266FB3A CRC64;

Query Match 40.0%; Score 6; DB 1; Length 558;

Best Local Similarity 100.0%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 10 OLAGR 15  
Db 289 OLAGR 294

## RESULT 7

Y079\_SCHPO STANDARD: PRT: 666 AA.  
ID Y079\_SCHPO  
AC O13670;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein C30B4.06 in chromosome II.  
DE SPCC30B4.06C OR P1079.  
GN Schizosaccharomyces pombe (Fission yeast).  
OS Schizosaccharomyces pombe.  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=20089027; PubMed=10620777;  
RA Machida M., Yamazaki S., Kunihito S., Tanaka T., Kushiya N., Jinno K.,  
RA Hatakeyama Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,  
RA Sakai M., Aoki K., Ogura K., Kudoh Y., Kikuchi H., Zhang M.Q.,  
RA Yanagida M.;  
RT "A 38 kb segment containing the cdc2 gene from the left arm of fission  
RT yeast chromosome II: sequence analysis and characterization of the  
RT genomic DNA and cDNAs encoded on the segment.";  
RL Yeast 16:71-80 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,  
RA Spouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gentles S., Goble A., Hamilton N., Harris D., Hidalgo J., Hodgson G.,  
RA Holtroft S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Voicakeert G., Aert R., Robben J., Grymoprez B.,  
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Botzys K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambolt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leclaire V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880 (2002).  
CC -1- SIMILARITY: BELONGS TO THE GIDA FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

```
CC -----
DR EMBL: AB004539; BAA21461.1; -
DR InterPro: IPR002218; GIDA.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01134; GIDA.1.
DR ProDom: PD003738; GIDA.1.
DR PROSITE: PS01280; GIDA.1.
DR PROSITE: PS01281; GIDA.2; 1.
DR Hypothetical protein.
KW SEQUENCE 666 AA; 74230 MW; 703D4529BDEBA83 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 1; Length 666;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 QLAAGR 15
|11111
DB 185 QLAAGR 190

RESULT 8
FTSK_COXBU STANDARD; PRT; 778 AA.
ID FTSK_COXBU
AC P39920;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cell division protein ftsk homolog.
GN FTSK OR SPOIIE.
OS Coccidia burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Coccidia group; Coccidia.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / Bratislava;
RX MEDLINE=94055499; PubMed=8237209;
RT "A sporulation gene in Coccidia burnetii?".
RL J. Vet. Med. B 40:366-370(1993).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE FTSK/SPOIIE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X75627; CAAS3289.1; -
DR InterPro: IPR002543; FtsK_SPOIIE.
DR Pfam: PF01580; FtsK_SPOIIE; 1.
KW Cell division; ATP-binding; Transmembrane.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 163 183 POTENTIAL.
FT NP_BIND 436 443 ATP (POTENTIAL).
SQ SEQUENCE 778 AA; 85169 MW; FCE6C15100422061 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 1; Length 778;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LQLAAG 14
|11111
DB 340 LQLAAG 345

RESULT 9
```

```
DP3A_NEIMA STANDARD; PRT; 1144 AA.
ID DP3A_NEIMA
AC Q93VX8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase III alpha subunit (EC 2.7.7.7).
GN DNAP OR NMA0632.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20722556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491."
RL Nature 404:502-506(2000).
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
+ [DNA](n).
CC -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
DIMERIZATION TO FORM THE POLIIT' COMPLEX. POLIIT' ASSOCIATES WITH
CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. DNAP
SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL162753; CAB83921.1; -
DR InterPro: IPR004013; PHP.C.
DR InterPro: IPR003141; PHP.N.
DR InterPro: IPR004805; PolC_alpha.
DR InterPro: IPR004365; tRNA_ant1.
DR Pfam: PF01336; tRNA_ant1; 1.
DR Pfam: PF02231; PHP_N; 1.
DR Pfam: PF02811; PHP_C; 1.
DR SMART: SM00481; POLIITac; 1.
DR TIGRFAMs: TIGR00594; polC; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW Complete proteome.
SQ SEQUENCE 1144 AA; 126975 MW; 95270191E29EB412 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 1; Length 1144;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RIPLOL 11
|11111
DB 1097 RIPLOL 1102

RESULT 10
DP3A_NEIMB STANDARD; PRT; 1144 AA.
ID DP3A_NEIMB
```

AC Q9JXZ2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DNA polymerase III alpha subunit (EC 2.7.7.7).  
 GN DNAE OR NMB1827.  
 OS Neisseria meningitidis (serogroup B).  
 OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.  
 NCBI\_TaxID=491.  
 RX SEQUENCE FROM N.A.  
 RC STRAIN-MC58 / Serogroup B;  
 RA MEDLINE=20175755; PubMed=10710307;  
 RA Tectelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
 RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,  
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
 RA Mason T., Clecko A., Parney D.S., Blair E., Citron H., Clark E.B.,  
 RA Cotton M.D., Ullendick T.R., Knout H., Qin H., Vamathevan J.,  
 RA Gill J., Scarlato V., Maignan V., Pizsa M., Grandi G., Sun L.,  
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappaport R., Venter J.C.;  
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain  
 MC58.";  
 RL Science 287:1809-1815(2000).  
 CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME  
 CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.  
 CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.  
 CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate  
 CC + [DNA](N).  
 CC -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA  
 CC CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE  
 CC DIMERIZATION TO FORM THE POLIIT' COMPLEX. POLIIT' ASSOCIATES WITH  
 CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,  
 CC AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. DNAE  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AE002532; AAF42162.1; -;  
 DR TIGR: NMB1827; -;  
 DR InterPro: IPR004013; PHP\_C.  
 DR InterPro: IPR003141; PHP\_N.  
 DR InterPro: IPR004805; POLC\_alpha.  
 DR InterPro: IPR004365; trna\_antl.  
 DR Pfam: PF01336; trna\_antl.1.  
 DR Pfam: PF02231; PHP\_N.1.  
 DR Pfam: PF02811; PHP\_C.1.  
 DR SMART: SM00481; POLIITAC.1.  
 DR TIGRFAMs: TIGRF00594; POLC.1.  
 KW Transferase: DNA-directed DNA polymerase; DNA replication;  
 KW Complete proteome.  
 SQ SEQUENCE 1144 AA; 127134 MW; 4CEABB86F90DD7EA CRC64;

Query Match 40.0%; Score 6; DB 1; Length 1144;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RLPQL 11  
 |||||  
 DB 1097 RLPQL 1102

RESULT 11  
 HOXY\_RHOOP

ID HOXY\_RHOOP STANDARD: PRT; 29 AA.  
 AC P22660;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE NAD-reducing hydrogenase hoxs delta subunit (EC 1.12.1.2) (Fragment).  
 GN HOXY.  
 OS Rhodococcus opacus (Nocardia opaca).  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
 NCBI\_TaxID=37919;  
 RX SEQUENCE.  
 RC STRAIN-1B;  
 RA MEDLINE=89231684; PubMed=2496982;  
 RA Zaborosch C., Schneider K., Schlegel H.G., Kratzin H.;  
 RT "Comparison of the NH2-terminal amino acid sequences of the four non-  
 RT identical subunits of the NAD-linked hydrogenases from Nocardia opaca  
 RT 1b and Alcaligenes eutrophus H16.";  
 RL Eur. J. Biochem. 181:175-180(1989).  
 CC -1- CATALYTIC ACTIVITY: H(2) + NAD(+) -> H(+) + NADH.  
 CC -1- COFACTOR: FMN, NICKEL, TWO 4FE-4S, A 3FE-4S, AND A 2FE-2S  
 CC CLUSTER.  
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA AND A GAMMA SUBUNITS (FLAVIN-  
 CC CONTAINING DIMER), AND A DELTA AND A NICKEL-CONTAINING BETA  
 CC SUBUNIT (HYDROGENASE DIMER).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: TO OTHER [NIFE] OR [NIFES] HYDROGENASES SMALL  
 CC SUBUNIT.  
 DR PIR: S03947; S03947.  
 KW Oxidoreductase; NAD; Iron-sulfur; 3fe-4s; Plasmid.  
 FT NON\_TER 29  
 SQ SEQUENCE 29 AA; 3159 MW; 1F54BC68F6AED6A CRC64;

Query Match 33.3%; Score 5; DB 1; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 LAAGR 15  
 |||||  
 DB 23 LAAGR 27

RESULT 12  
 REPL\_ECOLI  
 ID REPL\_ECOLI STANDARD: PRT; 93 AA.  
 AC P05830;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 01-OCT-1989 (Rel. 12, Last annotation update)  
 DE Regulatory protein repli (Fragment).  
 GN REPL.  
 OS Escherichia coli.  
 OG Plasmid R1162.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87040771; PubMed=2430262;  
 RA Kim K., Meyer R.J.;  
 RT "Copy-number of broad host-range plasmid R1162 is regulated by a  
 RT small RNA.";  
 RL Nucleic Acids Res. 14:8027-8046(1986).  
 CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN REGULATING THE PLASMID  
 CC COPY-NUMBER, INCREASING THE LEVEL OF THIS PROTEIN RESULTS IN A  
 CC HIGHER PLASMID COPY-NUMBER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04499; CAA28184.1; -.  
 DR PIR: A25346; RGECDI.  
 KW Plasmid; Plasmid copy control.  
 FT NON\_TER 93  
 SQ SEQUENCE 93 AA; 9413 MW; 0E767AAC26C7634A CRC64;  
 Query Match 33.3%; Score 5; DB 1; Length 93;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 9 LQIAA 13  
 Db 48 LQIAA 52  
 RESULT 13  
 ID FER1\_CVAPA STANDARD: PRT; 98 AA.  
 AC P17007; 98 AA.  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ferredoxin I.  
 OS PFMF.  
 OS Cyanophora paradoxa.  
 OS Cyanelle.  
 OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
 OX NCBI\_TaxID=2762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LB555 / Pringsheim;  
 RA Neumann-Spallart C., Brandtner M., Kraus M., Jakovitch J.,  
 RA Bayer M.G., Maier T.L., Schenk H.E.A., Loeffelhardt W.;  
 RT "The petf gene encoding ferredoxin I is located close to the str  
 RT operon on the cyanelle genome of Cyanophora paradoxa.";  
 RL FEBS Lett. 268:35-38(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91200662; PubMed=1901820;  
 RA Bryant D.A., Schluchter W.M., Stirewalt V.L.;  
 RT "Ferredoxin and ribosomal protein S10 are encoded on the cyanelle  
 RL genome of Cyanophora paradoxa.";  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LB555 / Pringsheim;  
 RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,  
 RA Bryant D.A.;  
 RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";  
 RL Plant Mol. Biol. Rep. 13:327-332(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LB555 / Pringsheim;  
 RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,  
 RA Farley J.Y., Schluchter W.M., Chung S., Neumann-Spallart C.,  
 RA Steiner J.M., Jakovitch J., Bohnert H.J., Bryant D.A.;  
 RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:  
 RT the genetic complexity of a primitive plastid.";  
 RL (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,  
 RL Schweinmler W. (eds.);  
 RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg  
 RL (1997).  
 RN [5]  
 RP SEQUENCE OF 1-12  
 RA Stevanovic S., Bayer M.G., Troger W., Schenk H.E.A.;  
 RT "Cyanophora paradoxa korsch.: ferredoxin partial amino-terminal  
 RT amino acid sequence, phylogenetic/taxonomic evidence.";  
 RL Endocyt. Cell Res. 6:219-226(1990).

CC -!- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER  
 CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.  
 CC -!- COFACTOR: BINDS 1 ZFE-2S CLUSTER.  
 CC -!- SUBCELLULAR LOCATION: CYANELLE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X52143; CAA36387.1; -.  
 DR EMBL: M35206; AAA31699.1; -.  
 DR EMBL: U30821; AAA81236.1; -.  
 DR PIR: S10427; FEKTL.  
 DR PIR: J00459; J00459.  
 DR HSSP: P00246; 4FXC.  
 DR InterPro: IPR000564; 2FE2S\_ferredoxin.  
 DR InterPro: IPR001041; Ferredoxin.  
 DR Pfam: PF00111; fer2; 1.  
 DR PRINTS: PR00159; 2FE2SFEROXIN.  
 DR PROSITE: PS00197; 2FE2S\_FERREDOXIN; 1.  
 KW Electron transport; Iron-sulfur; Cyanelle.  
 FT INIT\_MET 0  
 FT METAL 41 41 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 49 49 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 79 79 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 SQ SEQUENCE 98 AA; 10595 MW; 05061C34542FF77A CRC64;  
 Query Match 33.3%; Score 5; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 10 QIAAG 14  
 Db 70 QIAAG 74

RESULT 14  
 ID FER2\_APHSA STANDARD: PRT; 99 AA.  
 AC P00251; 99 AA.  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1988 (Rel. 09, Last annotation update)  
 DE Ferredoxin II.  
 OS Aphanothece sacrum.  
 OC Bacteria; Cyanobacteria; Chroococcales; Aphanothece.  
 OX NCBI\_TaxID=1122;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=78150873; PubMed=417074;  
 RA Hase T., Wakabayashi S., Wada K., Matsubara H.;  
 RT "Amino acid sequence of Aphanothece sacrum Ferredoxin II (minor  
 RT component). Structural characteristics and evolutionary  
 RT implications.";  
 RL J. Biochem. 83:761-770(1978).  
 CC -!- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER  
 CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.  
 CC -!- COFACTOR: BINDS 1 ZFE-2S CLUSTER.  
 CC -----  
 DR PIR: A00255; FEAH2.  
 DR HSSP: P00246; 4FXC.  
 DR InterPro: IPR000564; 2FE2S\_ferredoxin.  
 DR InterPro: IPR001041; Ferredoxin.  
 DR Pfam: PF00111; fer2; 1.  
 DR PRINTS: PR00159; 2FE2SFEROXIN.  
 DR PROSITE: PS00197; 2FE2S\_FERREDOXIN; 1.  
 KW Electron transport; Iron-sulfur.  
 FT METAL 41 41 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

```

FT METAL 49 49 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 80 80 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 99 AA: 10354 MW: 28D256A1CE86E00F CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 QLAG 14
   ||||
DB 71 QLAG 75

RESULT 15
YAIY ECOLI STANDARD: PRT: 102 AA.
ID YAIY ECOLI
AC P77669:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yaiy.
GN YAIY OR B0379 OR Z0475 OR ECS0429.
OS Escherichia coli, and
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schraum S., Davis R.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans F.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Ikeda T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [5]
RP SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000144; AAC73482.1; -.
DR EMBL: U73857; AAB18102.1; -.
DR EMBL: AE005217; AAC54725.1; -.
DR EMBL: AP002551; BAB3852.1; -.
DR EcoGene: EG14279; yaiy.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 75 95
SQ SEQUENCE 102 AA: 11424 MW: CA87D5A8E0F89F32 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 80;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWRHR 6
   ||||
DB 96 HWRHR 100

```

Search completed: November 27, 2002, 07:28:46  
 Job time : 5.44788 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:24:25 : Search time 4.2278 Seconds  
(Without alignments) 731.044 Million cell updates/sec

Title: US-09-893-615-2  
15

Perfect score: 15  
Sequence: 1 MHWNRHRIQLAAGR 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	46.7	160	09FTM6	Q9FTM6 Oryza sativ
2	7	46.7	168	0988K1	Q988K1 rhizobium 1
3	7	46.7	523	4 096DU2	Q96DU2 homo sapien
4	7	46.7	524	11 09R0H5	Q9R0H5 mus musculu
5	7	46.7	594	3 09PA43	Q9PA43 emericella
6	7	46.7	1549	3 09VP01	Q9VP01 drosophila
7	7	46.7	1869	16 092JV3	Q92JV3 rhizobium m
8	7	46.7	82	16 09KEX7	Q9KEX7 bacillus ha
9	6	40.0	111	16 09X2V3	Q9X2V3 escherichia
10	6	40.0	142	2 044529	Q44529 azotobacter
11	6	40.0	158	16 09CLH6	Q9CLH6 pasteurella
12	6	40.0	176	5 09VI03	Q9VI03 pasteurella
13	6	40.0	211	17 029897	Q29897 archaeoglob
14	6	40.0	218	2 P94955	P94955 micrococcus
15	6	40.0	231	15 056309	O56309 walleys epi
16	6	40.0	254	17 027483	O27483 methanobact

17	6	40.0	261	2 09KWS4	Q9KWS4 pseudomonas
18	6	40.0	262	17 08ZZA7	Q8ZZA7 pyrobaculum
19	6	40.0	266	2 056187	Q56187 streptomyc
20	6	40.0	272	11 09CS07	Q9CS07 mus musculu
21	6	40.0	283	11 08R3S2	Q8R3S2 mus musculu
22	6	40.0	291	5 09NVJ2	Q9NVJ2 haemochus
23	6	40.0	307	11 09EPG0	Q9EPG0 mus musculu
24	6	40.0	309	16 08YFH0	Q8YFH0 brucella me
25	6	40.0	317	16 09EBW3	Q9EBW3 streptomyc
26	6	40.0	330	16 09CC52	Q9CC52 mycobacteri
27	6	40.0	334	16 0920K3	Q920K3 rhizobium m
28	6	40.0	344	16 09RWF7	Q9RWF7 delinococcus
29	6	40.0	345	5 09VE32	Q9VE32 drosophila
30	6	40.0	347	11 054764	Q54764 mus musculu
31	6	40.0	358	13 057434	Q57434 fugu rubrip
32	6	40.0	387	11 09D845	Q9D845 mus musculu
33	6	40.0	395	12 092283	Q92283 molluscum c
34	6	40.0	413	6 028665	Q28665 oryctolagus
35	6	40.0	413	6 028666	Q28666 oryctolagus
36	6	40.0	413	6 007298	Q07298 oryctolagus
37	6	40.0	423	13 057433	Q57433 fugu rubrip
38	6	40.0	440	16 069661	Q69661 mycobacteri
39	6	40.0	449	10 09LIV0	Q9LIV0 oryza sativ
40	6	40.0	450	16 099027	Q99027 staphylococ
41	6	40.0	450	16 0931T9	Q931T9 staphylococ
42	6	40.0	456	6 062663	Q62663 oryctolagus
43	6	40.0	458	16 08RHY2	Q8RHY2 fusobacteri
44	6	40.0	467	2 09X4H6	Q9X4H6 streptomyc
45	6	40.0	475	16 08VIV5	Q8VIV5 mycobacteri

## ALIGNMENTS

## RESULT 1

Q9FTM6 PRELIMINARY; PRT; 160 AA.  
ID 09FTM6  
AC 09FTM6:  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE P0005A05.23 protein (P0482C06.6 protein).  
GN P0005A05.23 OR P0482C06.6.  
OS Oryza sativa (Rice)  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Euphorbiaceae; Oryzaeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
clone:P0005A05.23";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
clone:P0482C06.6";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP002863; BAB16919.1; -;  
DR EMBL; AP002845; BAB78610.1; -;  
SQ SEQUENCE 160 AA; 16198 MW; 86560C2B6AC8F777 CRC64;

Query Match 46.7%; Score 7; DB 10; Length 160;  
Best local similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 8 PQLAAG 14  
|||||  
DB 38 PQLAAG 44

## RESULT 2

0988K1

ID 0988K1 PRELIMINARY; PRT: 518 AA.

AC 0988K1: 01-OCT-2001 (TREMBLrel. 18, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical protein ml16708.

GN ML16708.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI\_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAFF303099;

RX MEDLINE=21082930; PubMed=11214968;

RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003009; BAB52949.1; -.

DR InterPro: IPR001633; EAL.

DR InterPro: IPR000160; GGDEF.

DR Pfam: PF00563; EAL; 1.

DR Pfam: PF00990; GGDEF; 1.

DR SMART: SM00267; DUF1; 1.

DR SMART: SM00052; DUF2; 1.

DR TIGRPMAS: TIGR00254; GGDEF; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 518 AA; 56502 MW; ED3561CD737C363E CRC64;

Query Match 46.7%; Score 7; DB 16; Length 518;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LQLAAGR 15

Db 142 LQLAAGR 148

## RESULT 3

096DU2

ID 096DU2 PRELIMINARY; PRT: 523 AA.

AC 096DU2: 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Keratin 6 lrs.

GN KRT6IRS.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=SCALP;

RA Rogers M.A., Langbein L., Winter H., Praetzel S., Ehmman C.,

RA Schweizer J.,

RT "K6irs, a new cytokeratin in the human hair follicle."

RL Submitt (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ308599; CAC43429.1; -.

DR InterPro: IPR001664; IF.

DR Pfam: PF00038; Filament; 1.

DR PROSITE: PS00226; IF; UNKNOWN\_1.

SQ SEQUENCE 523 AA; 57291 MW; 797F5655EE3A62D7 CRC64;

Query Match 46.7%; Score 7; DB 4; Length 523;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LQLAAGR 15

Db 330 LQLAAGR 336

## RESULT 4

09ROH5

ID 09ROH5 PRELIMINARY; PRT: 524 AA.

AC 09ROH5: 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Type II cytokeratin (Keratin protein K6irs).

GN KRT2-6G.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Aoki N., Tsujimoto T.;

RT "Isolation and characterization of a novel mouse type II

RT cytokeratin."

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6;

RA Porter R.M., McLean W.;

RT "Keratin K6irs is specific to the inner root sheath of hair follicles

RT in mouse and humans."

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1-SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

DR EMBL: AB033744; BA85657.1; -.

DR MGD: AY033497; AAK5110.1; -.

DR EMBL: MGI:1861586; Krt2-6g.

DR InterPro: IPR001664; IF.

DR InterPro: IPR002957; Keratin\_I.

DR InterPro: IPR003054; Keratin\_II.

DR Pfam: PF00038; Filament; 1.

DR PRINTS: PR01248; TYPE1KERATIN.

DR PRINTS: PR01276; TYPE2KERATIN.

DR PROSITE: PS00226; IF; 1.

KW Coiled coil; Intermediate filament.

SQ SEQUENCE 524 AA; 57382 MW; 45FF0CA4B440A72A CRC64;

Query Match 46.7%; Score 7; DB 11; Length 524;

Best Local Similarity 100.0%; Pred. No. 31;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LQLAAGR 15

Db 331 LQLAAGR 337

## RESULT 5

09P4A3

ID 09P4A3 PRELIMINARY; PRT: 594 AA.

AC 09P4A3: 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Camptothecin resistance conferring protein rcaA.

OS Emericella nidulans.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiiales; Trichocomaceae; Emericella.

OX NCBI\_TaxID=162425;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21258791; PubMed=11361337;



RA Bruschi G.C., de Souza C.C., Fagundes M.R., Dani M.A., Goldman M.H.,  
 RA Morris N.R., Liu L., Goldman G.H.:  
 RT "Sensitivity to camptothecin in *Aspergillus nidulans* identifies a  
 RT novel gene, *scatA*, related to the cellular DNA damage response."  
 RL Mol. Genet. Genomics 265:264-275(2001).  
 DR EMBL: AF228504; AAF81094.1; -  
 SO SEQUENCE 594 AA; 67513 MW; 19C4BE8E3DBD4505 CRC64;

Query Match 46.7%; Score 7; DB 3; Length 594;  
 Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RHRIPLQ 10  
 DB 489 RHRIPLQ 495

RESULT 6  
 Q9VP01 PRELIMINARY; PRT; 1549 AA.

DT 01-MAR-2000 (TRENBLREL. 13, Created)  
 DT 01-MAR-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)  
 DE CG11250 protein.  
 GN CG11250.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borovaya D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dushin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jajala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mallett B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weissbrock G.M., Weissbrock J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:  
 RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

DR EMBL: AE003595; AAF51761.1; -  
 DR FlyBase: FBgn0037113; CG11250.  
 DR InterPro: IPR002110; ANK\_PQZ.  
 DR InterPro: IPR002110; BTB\_PQZ.  
 DR InterPro: IPR01087; Lipase\_GDSL.  
 DR Pfam: PF00023; ank; 4.  
 DR Pfam: PF00651; BTB; 1.  
 DR SMART: SM00248; ANK; 2.  
 DR SMART: SM00225; BTB; 1.  
 DR PROSITE: PS50088; ANK\_REPEAT; 1.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE: PS50097; BTB; 1.  
 DR PROSITE: PS01098; LIPASE\_GDSL\_SER; UNKNOWN\_1.  
 KW ANK repeat; Repeat.  
 SO SEQUENCE 1549 AA; 166717 MW; ED837D9A83114DCF CRC64;

Query Match 46.7%; Score 7; DB 5; Length 1549;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PIQLANG 14  
 DB 1100 PIQLANG 1106

RESULT 7  
 Q92JY3 PRELIMINARY; PRT; 1869 AA.

DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Hypothetical signal peptide protein SMC03096.  
 GN R02997 OR SMC03096.  
 OS *Rhizobium meliloti* (*Sinorhizobium meliloti*).  
 OC *Rhizobium meliloti* (*Sinorhizobium meliloti*).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; *Sinorhizobium*.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barlier-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godard T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebaud P., Vandenberg M., Weidner S., Galibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT *Sinorhizobium meliloti* strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL: AL591792; CAC47576.1; -  
 DR InterPro: IPR000425; MIP\_family.  
 DR PROSITE: PS00221; MIP; UNKNOWN\_1.  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 1869 AA; 191305 MW; 32B51E6678708F9B CRC64;

Query Match 46.7%; Score 7; DB 16; Length 1869;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RIPLQIA 12  
 DB 813 RIPLQIA 819

RESULT 8  
 Q9KEX7 PRELIMINARY; PRT; 82 AA.  
 AC Q9KEX7;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Hypothetical protein BH0722.

GN BH0722.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL: AP001509; BAB04441.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 82 AA; 8895 MW; 2747633A66928A81 CRC64;

Query Match 40.0%; Score 6; DB 16; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LQLAAG 14  
 |||||  
 DB 18 LQLAAG 23

## RESULT 9

OBX279 PRELIMINARY; PRT; 111 AA.  
 ID OBX279;  
 AC OBX279;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein ECS4472.  
 GN ECS4472.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayaishi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shimagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12."  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL: AP002566; BAB37895.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 111 AA; 12340 MW; 9C49517DED123AD7 CRC64;

Query Match 40.0%; Score 6; DB 16; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PLQLAA 13  
 |||||  
 DB 89 PLQLAA 94

## RESULT 10

O44529 PRELIMINARY; PRT; 142 AA.  
 ID O44529;  
 AC O44529;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Potential molybdenum-pterin-binding-protein.

GN MODG.  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Azotobacter.  
 OX NCBI\_TaxID=354;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DW136;  
 RX MEDLINE=95394849; PubMed=7665518;  
 RA Mouncey N.J., Mitchenell L.A., Pau R.N.;  
 RT "Mutational analysis of genes of the mod locus involved in molybdenum  
 RT transport, homeostasis, and processing in Azotobacter vinelandii."  
 RL J. Bacteriol. 177:5294-5302(1995).  
 DR EMBL: Z49882; CAA90038.1; -;  
 DR InterPro: IPR004606; MOP.  
 DR InterPro: IPR005116; TOBE.  
 DR Pfam: PF03459; TOBE; 2.  
 DR TIGRfams: TIGR00638; MOP; 2.  
 SQ SEQUENCE 142 AA; 14359 MW; 514A7D97AEC2ED93 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LQLAAG 14  
 |||||  
 DB 47 LQLAAG 52

## RESULT 11

O9CLH6 PRELIMINARY; PRT; 158 AA.  
 ID O9CLH6  
 AC O9CLH6;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical protein PM1255.  
 GN PM1255.  
 OS Pasteurella multocida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Pasteurella.  
 OX NCBI\_TaxID=747;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PM70;  
 RX MEDLINE=21145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
 RT "Complete genomic sequence of Pasteurella multocida pm70."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 DR EMBL: AE006164; AAK03339.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 158 AA; 18065 MW; C9135DF60FAD7ADA CRC64;

Query Match 40.0%; Score 6; DB 16; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 QLAAGR 15  
 |||||  
 DB 32 QLAAGR 37

## RESULT 12

O9V103 PRELIMINARY; PRT; 176 AA.  
 ID O9V103  
 AC O9V103;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG10919 protein.  
 GN CG10919.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```

OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:
OC Ephydriidae: Drosophilidae: Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celmlker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.O., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Abmayant A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.A., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Meyers A.D., Dew I., Dietz S.M.,
RA Dodson K., Doop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA Relbert K., Remington K., Saunders R.D.C., Scheeler J., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupsky M.P., Smith T.,
RA Spletter E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003676; AAF54142.1; -;
DR Flybase: FBgn0037514; CG10919;
SQ SEQUENCE 176 AA; 21226 MW; D707C2448AAE5E53 CRC64;

Query Match 40.0%; Score 6; DB 5; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 HRIPLQ 10
DB 87 HRIPLQ 92

RESULT 13
OZ 029897
ID 029897 PRELIMINARY; PRT; 211 AA.
AC 029897;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Conserved hypothetical transmembrane protein.
GN AF0350.
OS Archaeoglobus fulgidus.
OC Archaea: Buryarchaeota: Archaeoglobi: Archaeoglobales:
OX Archaeoglobaceae: Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Claydon R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kurlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodok A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weisman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Galiani S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL: AE001080; AAB90884.1; -;
DR TIGR: AF0350; -;
DR InterPro: IPR002781; DUF81.
DR pfam: PF01925; DUF81.1.
KW Hypothetical protein; Transmembrane; Complete proteome.
SQ SEQUENCE 211 AA; 23303 MW; 927825BC1EFC1B46 CRC64;

Query Match 40.0%; Score 6; DB 17; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RPLQL 11
DB 37 RPLQL 42

RESULT 14
OZ 029897
ID 029897 PRELIMINARY; PRT; 218 AA.
AC 029897;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE H(+)-transporting ATP synthase (EC 3.6.1.34) (Fragment).
GN ATP H.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
OC Actinomycetales: Micrococcales: Micrococcaceae: Micrococcus.
OX NCBI_TaxID=1270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 4698;
RA Gruber G., Dose K., Nawroth T., Diel P., Cattaruzza M., Flach R.,
RT "Cloning and analysis of the atp-operon genes H and A from Micrococcus
RT luteus.";
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X84791; CAA59262.1; -;
DR InterPro: IPR000711; ATPsyn_OSCP.
DR pfam: PF00213; OSCP.1.
DR PRINTS: PR00125; ATPASEDELTA.
DR TIGRfams: TIGR01145; ATP_synth_delta.1.
KW Hydrolase.
KW NON_TER
FT
SQ SEQUENCE 218 AA; 22568 MW; F085D09081D93381 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LQLAAG 14
DB 166 LQLAAG 171

RESULT 15
OZ 056309
ID 056309 PRELIMINARY; PRT; 231 AA.
```

AC 056309;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Cyclin homolog.  
OS walleye epidermal hyperplasia virus' type 2.  
OC Viruses; Retrovird viruses; Retroviridae; Epsilonretrovirus.  
OX NCBI\_TaxID=64461;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lapierre L.A., Casey J.W., Holzschu D.L.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lapierre L.A., Holzschu D.L., Bowser P.R., Casey J.W.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY.  
DR EMBL: AF037569: AAC68475.1; -.  
DR EMBL: AF133052: AAC30056.1; -.  
DR InterPro:IPR004366; Cyclin.  
DR Pfam: PF00134; cyclin; 1.  
DR SMART: SM00385; CYCLIN; 1.  
KW Cell cycle; Cell division; Cyclin.  
SQ SEQUENCE 231 AA; 25367 MW; 95EA643029F27236 CRC64;  
  
Query Match 40.0%; Score 6; DB 15; Length 231;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
  
OY 9 LQLANG 14  
DB 64 LQLANG 69

Search completed: November 27, 2002, 07:30:24  
Job time : 10.4778 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 27, 2002, 07:26:06 : Search time 2.08494 Seconds

(Without alignments)  
211.682 Million cell updates/sec

Title: US-09-893-615-2

Sequence: 1 WHWHRIRPLQLAAGR 15

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	2	US-08-592-646A-28
2	15	100.0	15	2	US-08-592-646A-48
3	15	100.0	15	4	US-09-165-422-28
4	15	100.0	15	4	US-09-165-422-48
5	15	100.0	37	2	US-08-592-646A-59
6	15	100.0	37	4	US-09-165-422-59
7	14	93.3	15	2	US-08-592-646A-27
8	14	93.3	15	2	US-08-592-646A-29
9	14	93.3	15	4	US-09-165-422-27
10	14	93.3	15	4	US-09-165-422-29
11	13	86.7	15	2	US-08-592-646A-26
12	13	86.7	15	2	US-08-592-646A-30
13	13	86.7	15	4	US-09-165-422-26
14	13	86.7	15	4	US-09-165-422-30
15	12	80.0	15	2	US-08-592-646A-25
16	12	80.0	15	2	US-08-592-646A-31
17	12	80.0	15	4	US-09-165-422-25
18	12	80.0	15	4	US-09-165-422-31
19	11	73.3	11	2	US-08-592-646A-7
20	11	73.3	11	4	US-09-165-422-7
21	11	73.3	15	2	US-08-592-646A-9
22	11	73.3	15	2	US-08-592-646A-24
23	11	73.3	15	4	US-09-165-422-9
24	11	73.3	15	4	US-09-165-422-24
25	10	66.7	15	2	US-08-592-646A-23
26	10	66.7	15	2	US-08-592-646A-32
27	10	66.7	15	2	US-08-592-646A-40

28	10	66.7	15	4	US-09-165-422-23	Sequence 23, Appl
29	10	66.7	15	4	US-09-165-422-32	Sequence 32, Appl
30	10	66.7	15	4	US-09-165-422-40	Sequence 40, Appl
31	9	60.0	15	2	US-08-592-646A-22	Sequence 22, Appl
32	9	60.0	15	2	US-08-592-646A-33	Sequence 33, Appl
33	9	60.0	15	2	US-08-592-646A-41	Sequence 41, Appl
34	9	60.0	15	4	US-09-165-422-22	Sequence 22, Appl
35	9	60.0	15	4	US-09-165-422-33	Sequence 33, Appl
36	9	60.0	15	4	US-09-165-422-41	Sequence 41, Appl
37	8	53.3	15	2	US-08-592-646A-21	Sequence 21, Appl
38	8	53.3	15	2	US-08-592-646A-34	Sequence 34, Appl
39	8	53.3	15	2	US-08-592-646A-43	Sequence 43, Appl
40	8	53.3	15	4	US-09-165-422-21	Sequence 21, Appl
41	8	53.3	15	4	US-09-165-422-34	Sequence 34, Appl
42	8	53.3	15	4	US-09-165-422-43	Sequence 43, Appl
43	7	46.7	15	2	US-08-592-646A-20	Sequence 20, Appl
44	7	46.7	15	2	US-08-592-646A-35	Sequence 35, Appl
45	7	46.7	15	2	US-08-592-646A-42	Sequence 42, Appl

## ALIGNMENTS

RESULT 1  
US-08-592-646A-28  
: Sequence 28, Application US/08592646A  
: Patent No. 5851535  
: GENERAL INFORMATION:  
: APPLICANT: JOLIVET-REYNAUD, Colette  
: TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
: NUMBER OF SEQUENCES: 63  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSER: OLIFF & BERRIDGE, PLC  
: STREET: P.O. Box 19928  
: CITY: Alexandria  
: STATE: VA  
: COUNTRY: USA  
: ZIP: 22320  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/592,646A  
: FILING DATE: 26-JAN-1996  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Berridge, William P.  
: REGISTRATION NUMBER: 30,024  
: REFERENCE/DOCKET NUMBER: WPB 36923  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 703 836-6400  
: TELEFAX: 703 836-2787  
: INFORMATION FOR SEQ ID NO: 28:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 15 amino acids  
: TYPE: amino acid  
: TOPOLOGY: not relevant  
: MOLECULE TYPE: linear  
: US-08-592-646A-28

Query Match 100.0%; Score 15; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0;  
Oy 1 WHWHRIRPLQLAAGR 15  
Db 1 WHWHRIRPLQLAAGR 15

RESULT 2  
US-08-592-646A-48  
Sequence 48, Application US/08592646A  
Patent No. 5851535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIVE & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: MPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-48  
Query Match 100.0%; Score 15; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WHRRRIPQLAAGR 15  
DB 1 WHRRRIPQLAAGR 15  
RESULT 3  
US-09-165-422-28  
Sequence 28, Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIVE & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422

FILING DATE: 02-Oct-1998  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: MPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-165-422-28  
Query Match 100.0%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WHRRRIPQLAAGR 15  
DB 1 WHRRRIPQLAAGR 15  
RESULT 4  
US-09-165-422-48  
Sequence 48, Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIVE & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: MPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
US-09-165-422-48  
Query Match 100.0%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WHMRRIPQLAAGR 15  
| | | | | | | | | | |  
Db 1 WHMRRIPQLAAGR 15

## RESULT 5

US-08-592-646A-59  
; Sequence 59, Application US/08592646A  
; Patent No. 5851535  
; GENERAL INFORMATION:  
; APPLICANT: JOLIVET-REYNAUD, Colette  
; TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
; TITLE OF INVENTION: GONDII AND APPLICATIONS  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIEF & BERRIDGE, PLC  
; STREET: P.O. Box 19928  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/592,646A  
; FILING DATE: 26-JAN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berridge, William P.  
; REGISTRATION NUMBER: 30,024  
; REFERENCE/DOCKET NUMBER: WPB 36923  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703 836-6400  
; TELEFAX: 703 836-2787  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-592-646A-59  
; Query Match 100.0%; Score 15; DB 2; Length 37;  
; Best Local Similarity 100.0%; Pred. No. 4,7e-09;  
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WHMRRIPQLAAGR 15  
| | | | | | | | | | |  
Db 10 WHMRRIPQLAAGR 24

## RESULT 6

US-09-165-422-59  
; Sequence 59, Application US/09165422  
; Patent No. 6372897  
; GENERAL INFORMATION:  
; APPLICANT: JOLIVET-REYNAUD, Colette  
; TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
; TITLE OF INVENTION: GONDII AND APPLICATIONS  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIEF & BERRIDGE, PLC  
; STREET: P.O. Box 19928  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/165,422  
; FILING DATE: 02-Oct-1998  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berridge, William P.  
; REGISTRATION NUMBER: 30,024  
; REFERENCE/DOCKET NUMBER: WPB 36923  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703 836-6400  
; TELEFAX: 703 836-2787  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:  
US-09-165-422-59  
; Query Match 100.0%; Score 15; DB 4; Length 37;  
; Best Local Similarity 100.0%; Pred. No. 4,7e-09;  
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WHMRRIPQLAAGR 15  
| | | | | | | | | | |  
Db 10 WHMRRIPQLAAGR 24

## RESULT 7

US-08-592-646A-27  
; Sequence 27, Application US/08592646A  
; Patent No. 5851535  
; GENERAL INFORMATION:  
; APPLICANT: JOLIVET-REYNAUD, Colette  
; TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
; TITLE OF INVENTION: GONDII AND APPLICATIONS  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIEF & BERRIDGE, PLC  
; STREET: P.O. Box 19928  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/592,646A  
; FILING DATE: 26-JAN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berridge, William P.  
; REGISTRATION NUMBER: 30,024  
; REFERENCE/DOCKET NUMBER: WPB 36923  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703 836-6400  
; TELEFAX: 703 836-2787  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-592-646A-27

Query Match 93.3%; Score 14; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHRRHRIPLQLAG 14

DB 2 WHRRHRIPLQLAG 15

RESULT 8

US-08-592-646A-29  
Sequence 29, Application US/08592646A  
Patent No. 5851535

GENERAL INFORMATION:

APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MITOTOPIC POLYPEPTIDES OF TOXOPLASMA

TITLE OF INVENTION: GONDII AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF &amp; BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,646A

FILING DATE: 26-JAN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-592-646A-29

Query Match 93.3%; Score 14; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 2.2e-08;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 WHRRHRIPLQLAG 15

DB 1 WHRRHRIPLQLAG 14

RESULT 9

US-09-165-422-27  
Sequence 27, Application US/09165422  
Patent No. 6372897

GENERAL INFORMATION:

APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MITOTOPIC POLYPEPTIDES OF TOXOPLASMA

TITLE OF INVENTION: GONDII AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF &amp; BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/165,422

FILING DATE: 02-Oct-1998

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 27;

US-09-165-422-27

Query Match 93.3%; Score 14; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 2.2e-08;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHRRHRIPLQLAG 14

DB 2 WHRRHRIPLQLAG 15

RESULT 10

US-09-165-422-29  
Sequence 29, Application US/09165422  
Patent No. 6372897

GENERAL INFORMATION:

APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MITOTOPIC POLYPEPTIDES OF TOXOPLASMA

TITLE OF INVENTION: GONDII AND APPLICATIONS

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF &amp; BERRIDGE, PLC

STREET: P.O. Box 19928

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/165,422

FILING DATE: 02-Oct-1998

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36923

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 836-6400

TELEFAX: 703 836-2787

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:



LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 29;  
US-09-165-422-29

Query Match 93.3%; Score 13; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWRHRIPQLAAGR 15  
DB 1 HWRHRIPQLAAGR 14

RESULT 11  
US-08-592-646A-26  
Sequence 26 Application US/08592646A  
Patent No. 5651535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
TITLE OF INVENTION: GONDIT AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-30

Query Match 86.7%; Score 13; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WRRHRIPQLAAGR 13  
DB 1 WRRHRIPQLAAGR 13

RESULT 13  
US-09-165-422-26  
Sequence 26 Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
TITLE OF INVENTION: GONDIT AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-26

Query Match 86.7%; Score 13; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHRRHRIPQLAAG 13  
DB 3 WHRRHRIPQLAAG 15

RESULT 12  
US-08-592-646A-30  
Sequence 30 Application US/08592646A  
Patent No. 5651535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette

TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
TITLE OF INVENTION: GONDIT AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-30

Query Match 86.7%; Score 13; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WRRHRIPQLAAGR 13  
DB 1 WRRHRIPQLAAGR 13

RESULT 13  
US-09-165-422-26  
Sequence 26 Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
TITLE OF INVENTION: GONDIT AND APPLICATIONS  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-592-646A-26

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
MOLECULE TYPE: peptide  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-165-422-26

Query Match 86.7%; Score 13; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMHRIPQLAA 13  
Db 3 WHMHRIPQLAA 15

RESULT 14  
US-09-165-422-30  
Sequence 30; Application US/09165422  
Patent No. 6372897  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/165,422  
FILING DATE: 02-Oct-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
MOLECULE TYPE: peptide  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
US-09-165-422-30

Query Match 86.7%; Score 13; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WHMHRIPQLAAGR 15  
Db 1 WHMHRIPQLAAGR 13

RESULT 15  
US-08-592-646A-25  
Sequence 25; Application US/08592646A  
Patent No. 5851535  
GENERAL INFORMATION:  
APPLICANT: JOLIVET-REYNAUD, Colette  
TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,646A  
FILING DATE: 26-Jan-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Berridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 36923  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 836-6400  
TELEFAX: 703 836-2787  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
MOLECULE TYPE: peptide  
US-08-592-646A-25

Query Match 80.0%; Score 12; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WHMHRIPQLA 12  
Db 4 WHMHRIPQLA 15

Search completed: November 27, 2002, 07:32:09  
Job time: 3.08494 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:28:06 : Search time 1.38996 Seconds  
(without alignments)  
171.849 Million cell updates/sec

Title: US-09-893-615-2  
Perfect score: 15  
Sequence: 1 WHMRHRIPLQLAGR 15

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 102317 seqs, 15924203 residues

Word size: 0

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	10	US-09-893-615-2
2	15	100.0	15	10	US-09-893-615-2
3	15	100.0	15	10	US-09-893-615-9
4	15	100.0	15	10	US-09-893-615-17
5	15	100.0	15	10	US-09-893-615-19
6	15	100.0	15	10	US-09-893-615-41
7	14	93.3	19	10	US-09-893-615-15
8	6	40.0	228	10	US-09-893-737-142
9	6	40.0	290	10	US-09-815-242-12356
10	6	40.0	411	10	US-09-815-242-5746
11	6	40.0	585	10	US-09-815-242-11923
12	6	40.0	741	9	US-09-891-762-2
13	5	33.3	29	10	US-09-982-172-16
14	5	33.3	29	10	US-09-982-172-215
15	5	33.3	41	10	US-09-864-761-45173
16	5	33.3	46	10	US-09-864-761-46783
17	5	33.3	50	10	US-09-864-761-48053
18	5	33.3	54	10	US-09-925-297-888
19	5	33.3	57	10	US-09-864-761-48156

20	5	33.3	62	10	US-09-864-761-33340	Sequence 33340, A
21	5	33.3	80	10	US-09-864-761-48387	Sequence 48387, A
22	5	33.3	83	10	US-09-726-643-84	Sequence 84, Appl
23	5	33.3	88	8	US-08-424-550B-275	Sequence 275, Appl
24	5	33.3	92	10	US-09-867-550-494	Sequence 494, Appl
25	5	33.3	97	10	US-09-864-761-43414	Sequence 43414, A
26	5	33.3	97	12	US-10-036-342-16	Sequence 16, Appl
27	5	33.3	97	12	US-10-052-586-468	Sequence 468, Appl
28	5	33.3	111	10	US-09-764-898-295	Sequence 295, Appl
29	5	33.3	133	10	US-09-867-550-372	Sequence 372, Appl
30	5	33.3	140	10	US-09-864-761-43849	Sequence 43849, A
31	5	33.3	155	10	US-09-734-569-118	Sequence 118, Appl
32	5	33.3	180	10	US-09-771-730-131	Sequence 131, Appl
33	5	33.3	184	10	US-09-764-864-837	Sequence 827, Appl
34	5	33.3	190	9	US-09-895-913A-132	Sequence 132, Appl
35	5	33.3	193	10	US-09-941-936A-2	Sequence 2, Appl
36	5	33.3	202	10	US-09-811-284-151	Sequence 151, Appl
37	5	33.3	203	10	US-09-912-020-338	Sequence 338, Appl
38	5	33.3	215	10	US-09-924-256A-36	Sequence 36, Appl
39	5	33.3	216	10	US-09-771-730-50	Sequence 50, Appl
40	5	33.3	216	10	US-09-771-730-143	Sequence 143, Appl
41	5	33.3	216	10	US-09-771-730-146	Sequence 146, Appl
42	5	33.3	216	10	US-09-747-155-427	Sequence 427, Appl
43	5	33.3	218	12	US-10-011-033-2	Sequence 2, Appl
44	5	33.3	218	12	US-10-011-033-6	Sequence 6, Appl
45	5	33.3	218	12	US-10-011-033-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-09-893-615-2  
Sequence 2, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Flischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elinaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: Linear

```
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-893-615-2

Query Match
Best Local Similarity 100.0%; Score 15; DB 10; Length 15;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHMRHRIPQLAAGR 15
Db 1 WHMRHRIPQLAAGR 15

RESULT 2
US-09-893-615-9
Sequence 9, Application US/09893615
Patent No. US20020082395A1
GENERAL INFORMATION:
APPLICANT: FISCHER, Gerald W.
Schuman, Richard F.
Wong, Hing
Stinson, Jeffrey L.
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM
POSITIVE BACTERIA
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/893,615
FILING DATE: 29-Jun-2001
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 04995, 0041-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-893-615-9

Query Match
Best Local Similarity 100.0%; Score 15; DB 10; Length 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHMRHRIPQLAAGR 15
Db 3 WHMRHRIPQLAAGR 17

RESULT 3
US-09-893-615-17
Sequence 17, Application US/09893615
Patent No. US20020082395A1
GENERAL INFORMATION:
APPLICANT: FISCHER, Gerald W.
Schuman, Richard F.
Wong, Hing
Stinson, Jeffrey L.
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM
POSITIVE BACTERIA
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/893,615
FILING DATE: 29-Jun-2001
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 04995, 0041-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-893-615-17

Query Match
Best Local Similarity 100.0%; Score 15; DB 10; Length 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHMRHRIPQLAAGR 15
Db 3 WHMRHRIPQLAAGR 17

RESULT 4
US-09-893-615-19
Sequence 19, Application US/09893615
Patent No. US20020082395A1
GENERAL INFORMATION:
APPLICANT: FISCHER, Gerald W.
Schuman, Richard F.
Wong, Hing
Stinson, Jeffrey L.
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM
POSITIVE BACTERIA
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/893,615
FILING DATE: 29-Jun-2001
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 04995, 0041-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-893-615-19
```

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-893-615-19

Query Match 100.0%; Score 15; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WHMRRIPQLAAGR 15  
|||||  
Db 3 WHMRRIPQLAAGR 17

RESULT 5  
US-09-893-615-41  
Sequence 41, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 41:  
US-09-893-615-41

Query Match 100.0%; Score 15; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WHMRRIPQLAAGR 15  
|||||  
Db 3 WHMRRIPQLAAGR 17

RESULT 6  
US-09-893-615-57  
Sequence 57, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P.  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 57:  
US-09-893-615-57

Query Match 100.0%; Score 15; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WHMRRIPQLAAGR 15  
|||||  
Db 3 WHMRRIPQLAAGR 17

RESULT 7  
US-09-893-615-15  
Sequence 15, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.

TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOPEPTIC ACID OF GRAM  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
DUNNER, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/893,615  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Elinaudi, Carol P.  
REGISTRATION NUMBER: 37,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4400  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-893-615-15  
Query Match 93.3%; Score 14; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 HMRHRIPLQLAGR 15  
DB 4 HMRHRIPLQLAGR 17  
RESULT 8  
US-09-893-737-142  
Sequence 142, Application US/09893737  
Patent No. US20020110855A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Presnell, Scott R.  
TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS  
FILE REFERENCE: 00-41  
CURRENT APPLICATION NUMBER: US/09/893,737  
CURRENT FILING DATE: 2001-06-28  
PRIOR APPLICATION NUMBER: US 60/215,446  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 329  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 142  
LENGTH: 228  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-893-737-142  
Query Match 40.0%; Score 6; DB 10; Length 228;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 10 QLAAGR 15  
DB 200 QLAAGR 205

RESULT 9  
US-09-815-242-12356  
Sequence 12356, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12356  
LENGTH: 290  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12356  
Query Match 40.0%; Score 6; DB 10; Length 290;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 7 IPLQLA 12  
DB 241 IPLQLA 246  
RESULT 10  
US-09-815-242-5746  
Sequence 5746, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727 ;  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 5746  
LENGTH: 411  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5746

Query Match 40.0%; Score 6; DB 10; Length 411;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IPLOLA 12  
|||||  
DB 371 IPLOLA 376

RESULT 11  
US-09-815-242-11923  
Sequence 11923, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyckind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 11923  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-815-242-11923

Query Match 40.0%; Score 6; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 9 IOLAAG 14  
|||||

DB 142 IOLAAG 147

RESULT 12  
US-09-891-762-2  
Sequence 2, Application US/09891762  
Patent No. US20020156253A1  
GENERAL INFORMATION:  
APPLICANT: Curtis, Rory A.J.  
APPLICANT: Silos-Santiago, Immaculada  
TITLE OF INVENTION: 48000 AND 52920, NOVEL HUMAN CALCIUM  
CHANNELS AND USES THEREOF  
FILE REFERENCE: MNI-170  
CURRENT APPLICATION NUMBER: US/09/891,762  
CURRENT FILING DATE: 2001-06-26  
PRIOR APPLICATION NUMBER: US 60/214,176  
PRIOR FILING DATE: 2000-06-26  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 741  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-891-762-2

Query Match 40.0%; Score 6; DB 9; Length 741;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PLOLA 13  
|||||  
DB 344 PLOLA 349

RESULT 13  
US-09-982-172-16  
Sequence 16, Application US/09982172  
Patent No. US20020137119A1  
GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBIO  
TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE  
FILE REFERENCE: 01/22283  
CURRENT APPLICATION NUMBER: US/09/982,172  
CURRENT FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 16  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-16

Query Match 33.3%; Score 5; DB 10; Length 29;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LAAGR 15  
|||||  
DB 25 LAAGR 29

RESULT 14  
US-09-982-172-215  
Sequence 215, Application US/09982172  
Patent No. US20020137119A1  
GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBIO  
TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE

```

; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 215
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-215

Query Match.          33.3%; Score 5; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LAAGR 15
DB 25 LAAGR 29

RESULT 15
US-09-864-761-45173
; Sequence 45173; Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45173
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157392.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.52
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.52
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.51
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.48
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.46
; OTHER INFORMATION: EST_HUMAN HIT: BF224452.1, EVALU6 7.00e-17
; OTHER INFORMATION: SWISSPROT HIT: P25109, EVALU6 8.30e+00
US-09-864-761-45173

Query Match          33.3%; Score 5; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IPLQL 11
DB 24 IPLQL 28

Search completed: November 27, 2002, 07:32:52
Job time : 1.38996 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 04:43:17 ; Search time 1454.52 Seconds

(without alignments)  
7383.171 Million cell updates/sec

Title: US-09-893-615-86

Perfect score: 369

Sequence: 1 GAAGTGAATGCTGCTGAGACTC.....CCTCAGTCACCGCTCTCTCA 369

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched:

2054640 seqs, 14551402878 residues

Word size :

0

Total number of hits satisfying chosen parameters:

4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pal:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pt:\*

10: gb\_ro:\*

11: gb\_sls:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sls:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rtd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	145	39.3	316	10	AF118969	AF118969 Mus muscu
2	145	39.3	366	10	AF178591	AF178591 Mus muscu
3	145	39.3	366	10	AF178595	AF178595 Mus muscu
4	145	39.3	366	10	AF178596	AF178596 Mus muscu
5	144	39.0	312	10	MUSIGHACT	M34735 Mouse Ig H-
6	144	39.0	337	10	MMU5546	U55546 Mus musculu
7	144	39.0	357	10	MUSIGHXT	M21470 Mouse Ig ac
8	144	39.0	360	10	MUSIGHVRA	D50376 Mouse mRNA
9	144	39.0	474	10	MUSIGHCLT	M20829 Mouse IgG2a
10	144	39.0	661	10	AF064446	AF064446 Mus muscu
11	143	38.8	300	10	MMIGH1	X59104 M. musculus
12	143	38.8	663	10	AF064445	AF064445 Mus muscu
13	141	38.2	299	10	AF072804	AF072804 Mus muscu
14	138	37.4	360	10	MDIGVAVQ	Z22080 M. domestlicu
15	132	35.8	242	10	AF023222	AF023222 Mus muscu
16	132	35.8	330	6	AR108680	AR108680 Sequence
17	131	35.5	247	10	AF023223	AF023223 Mus muscu
18	130	35.2	345	6	AR108686	AR108686 Sequence
19	116	31.4	315	10	MMU55466	U55466 Mus musculu
20	112	30.4	353	10	AF321939	AF321939 Mus muscu
21	107	29.0	421	10	MMIGH1HV	X78107 M. musculus
22	98	26.6	375	10	S63182	S63182 Ig VH-antli-
23	93	25.2	301	10	MMU55465	U55465 Mus musculu
24	93	25.2	337	10	MMU55545	U55545 Mus musculu
25	87	23.6	348	10	AF468835	AF468835 Mus muscu
26	86	22.3	360	10	MDIGVAVR	Z22082 M. domestlicu
27	84	22.8	332	10	MMU55520	U55520 Mus musculu
28	80	21.7	357	6	I73737	I73737 Sequence 1
29	80	21.7	1797	6	I73745	I73745 Sequence 17
30	76	20.6	304	10	S40679	S40679 ADI heavy c
31	76	20.6	305	10	AF072799	AF072799 Mus muscu
32	76	20.6	351	10	AF006586	AF006586 Mus muscu
33	76	20.6	475	10	MUSIGHAAAC	M60022 Mouse Ig he
34	76	20.6	663	10	AF064443	AF064443 Mus muscu
35	76	20.6	185591	2	AC073561	AC073561 Mus muscu
36	74	20.1	360	10	AF006574	AF006574 Mus muscu
37	72	19.5	333	10	MMU55544	U55544 Mus musculu
38	69	18.7	322	10	AF118987	AF118987 Mus muscu
39	69	18.7	663	10	AF064444	AF064444 Mus muscu
40	68	18.4	231	10	S60842	S60842 AD2 heavy c
41	68	18.4	351	10	AF030229	AF030229 Mus muscu
42	64	17.3	366	10	MUSIGH339A	L30140 Mus musculu
43	63	17.1	357	10	MUSIGHUHF	M59948 Mouse Ig ac
44	63	17.1	360	10	MDIGVAVT	Z22093 M. domestlicu
45	61	16.5	251	10	AF133813	AF133813 Mus muscu

#### ALIGNMENTS

RESULT 1  
AF118969  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AF118969 316 bp mRNA linear ROD 22-APR-1999  
Mus musculus  
Mus musculus  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 316)  
Seidl,R.J., Wilschre,J.A., Mackenzie,J.D., Kantor,A.B.,  
Herzenberg,L.A. and Herzenberg,L.A.

TITLE Predominant VH genes expressed in innate antibodies are associated with distinctive antigen-binding sites  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (5), 2262-2267 (1999)  
MEDLINE 99162592  
PUBMED 10051629  
REFERENCE 2 (bases 1 to 316)  
AUTHORS Seldi, K.J., Milshire, J.A., Mackenzie, J.D., Herzenberg, L.A., Herzenberg, L.A. and Kantor, A.B.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUN-1999) Department of Genetics, Stanford University School of Medicine, Beckman Center, B007, Stanford, CA 94305-5125, USA

FEATURES  
source Location/Qualifiers  
1. 316  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="AE0203"  
<1..>316  
/codon\_start=-2  
/product="Immunoglobulin heavy chain VH10-JH4 region"  
/protein\_id="A026740.1"  
/db\_xref="GI:4633368"  
/translation="GGLVOPKSLKSCAASGFENTYAMNWRQAPGKLEWVARISKSNNVATRYADSVKDRFTISRDSQSMLYLQNNLKTEDTAMTCVROGRTGNAMD YWC"

CDS  
BASE COUNT 95 a 65 c 79 g 77 t  
ORIGIN

Query Match 39.3%; Score 145; DB 10; Length 316;  
Best Local Similarity 99.5%; Pred. No. 3.5e-77;  
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 TGGGTCCGCCAGGCTCCAGAAAGGTTTGGATGCTGCTCCCATTAAGAATAAAAGT 165  
|||||  
Db 83 TGGGTCCGCCAGGCTCCAGAAAGGTTTGGATGCTGCTCCCATTAAGAATAAAAGT 142  
|||||

QY 166 AATAATATATGCAACATTTTATGCCGATTCACTGAAGACAGCTTCCAGAGAT 225  
|||||  
Db 143 AATAATATATGCAACATTTTATGCCGATTCACTGAAGACAGCTTCCAGAGAT 202  
|||||

QY 226 GATTACACAAACATGCTCTATCTCAATGAACAACCTGAAGACAGACAGCCATG 285  
|||||  
Db 203 GATTACACAAACATGCTCTATCTCAATGAACAACCTGAAGACAGACAGCCATG 262  
|||||

QY 286 TATTACTGTGTGAGAC 301  
|||||  
Db 263 TATTACTGTGTGAGAC 278  
|||||

RESULT 2  
AF178591 366 bp mRNA linear ROD 22-MAY-2000  
LOCUS Mus musculus 23-7 Immunoglobulin heavy chain variable region mRNA,  
DEFINITION partial cds.  
ACCESSION AF178591  
VERSION AF178591.1 GI:5853171  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 366)  
AUTHORS Puterman, C., Deocharan, B. and Diamond, B.  
TITLE Molecular analysis of the autoantibody response in peptide-induced autoimmunity  
JOURNAL J. Immunol. 164 (5), 2542-2549 (2000)  
MEDLINE 20143847  
PUBMED 10679092  
REFERENCE 2 (bases 1 to 366)  
AUTHORS Puterman, C., Deocharan, B. and Diamond, B.  
TITLE Direct Submission  
JOURNAL Submitted (17-AUG-1999) Rheumatology/Medicine, Albert Einstein

FEATURES  
source Location/Qualifiers  
1. 366  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="23-7"  
/cell\_type="hybridoma"  
/note="from peptide-immunized mouse"  
<1..>366  
/note="anti-peptide antibody"  
/codon\_start=-1  
/product="Immunoglobulin heavy chain variable region"  
/protein\_id="A054347.1"  
/db\_xref="GI:5853172"  
/translation="DVKVVEGGGLVOPKSLKSCAASGFENTYAMNWRQAPGKLEWVARIRSKSNVATRYADSVKDRFTISRDSQSMLYLQNNLKTEDTAMTCVRRSGSRKYVDVWAGTYYVSS"

CDS  
BASE COUNT 101 a 84 c 95 g 86 t  
ORIGIN

Query Match 39.3%; Score 145; DB 10; Length 366;  
Best Local Similarity 99.5%; Pred. No. 3.6e-77;  
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 TGGGTCCGCCAGGCTCCAGAAAGGTTTGAATGCTGCTCCCATTAAGAATAAAAGT 165  
|||||  
Db 106 TGGGTCCGCCAGGCTCCAGAAAGGTTTGAATGCTGCTCCCATTAAGAATAAAAGT 165  
|||||

QY 166 AATAATATATGCAACATTTTATGCCGATTCACTGAAGACAGCTTCCAGAGAT 225  
|||||  
Db 166 AATAATATATGCAACATTTTATGCCGATTCACTGAAGACAGCTTCCAGAGAT 225  
|||||

QY 226 GATTACACAAACATGCTCTATCTCAATGAACAACCTGAAGACAGACAGCCATG 285  
|||||  
Db 226 GATTACACAAACATGCTCTATCTCAATGAACAACCTGAAGACAGACAGCCATG 285  
|||||

QY 286 TATTACTGTGTGAGAC 301  
|||||  
Db 286 TATTACTGTGTGAGAC 301  
|||||

RESULT 3  
AF178595 366 bp mRNA linear ROD 22-MAY-2000  
LOCUS Mus musculus 8-3 Immunoglobulin heavy chain variable region mRNA,  
DEFINITION partial cds.  
ACCESSION AF178595  
VERSION AF178595.1 GI:5853179  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 366)  
AUTHORS Puterman, C., Deocharan, B. and Diamond, B.  
TITLE Molecular analysis of the autoantibody response in peptide-induced autoimmunity  
JOURNAL J. Immunol. 164 (5), 2542-2549 (2000)  
MEDLINE 20143847  
PUBMED 10679092  
REFERENCE 2 (bases 1 to 366)  
AUTHORS Puterman, C., Deocharan, B. and Diamond, B.  
TITLE Direct Submission  
JOURNAL Submitted (17-AUG-1999) Rheumatology/Medicine, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA  
FEATURES  
source Location/Qualifiers  
1. 366  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="8-3"  
/cell\_type="hybridoma"

```
/note="from peptide-immunized mouse"
<1..>366
/codon_start=1
/note="anti-peptide antibody"
/product="immunoglobulin heavy chain variable region"
/db_xref="GI:5853180"
/translation="DVRVYSGGGIYVQKSGIKLSCASGFTFNTYANMWRQAPGK
LEWVARIRKSNNTATYADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYCVRHG
LGRWYFDWVGAGTIVTSS"
BASE COUNT      99 a      82 c      99 g      86 t
ORIGIN
Query Match      39.3%; Score 145; DB 10; Length 366;
Best Local Similarity 99.5%; Pred. No. 3.6e-77;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 165
|||||
Db 106 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 165
|||||

QY 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 225
|||||
Db 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 225
|||||

QY 226 GATTCAAAAGCATGCTCTATCTGCAAAATGAACACTTGAAACTGAGACACAGCCATG 285
|||||
Db 226 GATTCAAAAGCATGCTCTATCTGCAAAATGAACACTTGAAACTGAGACACAGCCATG 285
|||||

QY 286 TATTACTGTGTGAGAC 301
|||||
Db 286 TATTACTGTGTGAGAC 301
|||||

RESULT 4
AF178596      366 bp      mRNA      linear      ROD 22-MAY-2000
LOCUS
DEFINITION Mus musculus 39-9 immunoglobulin heavy chain variable region mRNA,
partial cds.
ACCESSION AF178596
VERSION AF178596.1 GI:5853181
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 366)
AUTHORS Puterman,C., Deocharan,B. and Diamond,B.
TITLE Mammalia; Eutheria; Rodentia; Sciurognathl; Muridae; Mus.
Molecular analysis of the autoantibody response in peptide-induced
autoimmunity
JOURNAL J. Immunol. 164 (5), 2542-2549 (2000)
MEDLINE 20143847
PUBMED 10679092
REFERENCE 2 (bases 1 to 366)
AUTHORS Puterman,C., Deocharan,B. and Diamond,B.
TITLE Mammalia; Eutheria; Rodentia; Sciurognathl; Muridae; Mus.
Direct Submission
JOURNAL Submitted (17-AUG-1999) Rheumatology/Medicine, Albert Einstein
College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA
FEATURES
source
1..366
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/cell_line="39-9"
/cell_type="hybridoma"
/note="from peptide-immunized mouse"
<1..>366
/note="anti-peptide antibody"
/codon_start=1
/product="immunoglobulin heavy chain variable region"
/protein_id="AAD54352.1"
/db_xref="GI:5853182"
/translation="DVKYVESGGGLVQPKSLKLSCAASGFTFNTYANMWRQAPGK

BASE COUNT      95 a      69 c      73 g      75 t
ORIGIN
Query Match      39.0%; Score 144; DB 10; Length 312;
Best Local Similarity 99.5%; Pred. No. 1.5e-76;
Matches 194; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 165
|||||
Db 68 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 127
|||||

QY 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 225
|||||
Db 128 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 187
|||||

/note="from peptide-immunized mouse"
<1..>366
/codon_start=1
/note="anti-peptide antibody"
/product="immunoglobulin heavy chain variable region"
/protein_id="AAD54352.1"
/db_xref="GI:5853182"
/translation="DVKYVESGGGLVQPKSLKLSCAASGFTFNTYANMWRQAPGK

BASE COUNT      100 a      82 c      98 g      86 t
ORIGIN
Query Match      39.3%; Score 145; DB 10; Length 366;
Best Local Similarity 99.5%; Pred. No. 3.6e-77;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 165
|||||
Db 106 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 165
|||||

QY 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 225
|||||
Db 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 225
|||||

QY 226 GATTCAAAAGCATGCTCTATCTGCAAAATGAACACTTGAAACTGAGACACAGCCATG 285
|||||
Db 226 GATTCAAAAGCATGCTCTATCTGCAAAATGAACACTTGAAACTGAGACACAGCCATG 285
|||||

QY 286 TATTACTGTGTGAGAC 301
|||||
Db 286 TATTACTGTGTGAGAC 301
|||||

RESULT 5
MUSIGHACT      312 bp      mRNA      linear      ROD 27-APR-1993
LOCUS
DEFINITION Mouse Ig H-chain mRNA VDJ-region, partial cds.
ACCESSION M34735
VERSION M34735.1 GI:194746
KEYWORDS D-region; J-region; V-region; immunoglobulin heavy chain; processed
gene.
SOURCE Mouse (strain BALB/c) adult spleen anti-HA hybridoma H220-3, CDNA
to mRNA.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 312)
AUTHORS Clarke,S.H., Staudt,L.M., Kaveler,J., Schwartz,D., Gerhard,W.U. and
Weigert,M.G.
TITLE V region gene usage and somatic mutation in the primary and
secondary responses to influenza virus hemagglutinin
JOURNAL J. Immunol. 144 (7), 2795-2801 (1990)
MEDLINE 90203619
PUBMED 2108213
FEATURES
source
1..312
/organism="Mus musculus"
/db_xref="taxon:10090"
<1..>312
/note="Ig H-chain VDJ-region"
/codon_start=2
/protein_id="AAA38022.1"
/db_xref="GI:194747"
/translation="FKGSLKLSCAASGFTFNTYANMWRQAPGKLEWVARIRKSNNTATYADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYCVRAAYAMDYWGQCTSY
TV"

BASE COUNT      95 a      69 c      73 g      75 t
ORIGIN
Query Match      39.0%; Score 144; DB 10; Length 312;
Best Local Similarity 99.5%; Pred. No. 1.5e-76;
Matches 194; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 165
|||||
Db 68 TGGGTCGGCAGGCTCCAGGAAGGTTGGATGGTTCGTCGATTAAGATAAAGT 127
|||||

QY 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 225
|||||
Db 128 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGTTCAACCATCTCCAGAGAT 187
|||||
```

QY 226 GATTACAAACGATGCTCTATCTGCAATGAACAACCTTGAAGAGACACAGCCATG 285  
|||||  
Db 188 GATTACAAACGATGCTCTATCTGCAATGAACAACCTTGAAGAGACACAGCCATG 247  
|||||  
QY 286 TATTACTGTGTGAGA 300  
|||||  
Db 248 TATTACTGTGTGAGA 262  
|||||

RESULT 6  
LOCUS MMU55546 337 bp mRNA linear ROD 14-SEP-2001  
DEFINITION Mus musculus anti-DNA immunoglobulin heavy chain IgG mRNA, antibody  
U55546  
VERSION U55546  
KEYWORDS U55546.1 GI:1872412  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 337)  
AUTHORS Krishnan,M.R., Jou,N.T. and Marion,T.N.  
TITLE Correlation between the amino acid position of arginine in VH-CDR3  
and specificity for native DNA among autoimmune antibodies  
JOURNAL J. Immunol. 157 (6), 2430-2439 (1996)  
MEDLINE 9639071  
PUBMED 8805642  
REFERENCE 2 (bases 1 to 337)  
AUTHORS Marion,T.N.  
TITLE Direct Submission  
JOURNAL Submitted (18-APR-1996) Dept: of Microbiology/Immunology,  
University of Tennessee, 858 Madison Ave., Memphis, TN 38163, USA  
FEATURES  
SOURCE location/Qualifiers.  
1..337  
/organism="Mus musculus"  
/strain="(NZB x NZW) F1"  
/db\_xref="taxon:10090"  
/cell\_line="452P.108"  
/cell\_type="hybridoma"  
/note="mouse number 452; clonally related to hybridomas  
452P.119 and 452P.84"  
<1..>337  
/codon\_start=2  
/product="anti-DNA immunoglobulin heavy chain IgG"  
/protein\_id="AAB49145.1"  
/db\_xref="GI:1872413"  
/translation="VQKESLKLSCAAGSFSPNTNANMVRQAPKGLKLEWVARIRSKS  
NNATYVADSVKDRFTISRDSOSMLYLNKLTEDTAMYCVREGSYSLYWXDVW  
GAGTIVVSS"  
V\_region  
BASE COUNT 96 a 75 c 82 g 80 t 4 others  
ORIGIN  
Query Match 39.0%; Score 144; DB 10; Length 337;  
Best Local Similarity 99.5%; Pred. No. 1.5e-76;  
Matches 194; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 254 TATTACTGTGTGAGA 268  
|||||

RESULT 7  
LOCUS MUSIGHXT 357 bp mRNA linear ROD 27-APR-1993  
DEFINITION Mouse Ig active H chain mRNA V-region, MRL-DNA4.  
ACCESSION M21470.1 GI:196220  
VERSION M21470.1  
KEYWORDS C-region; V-region; immunoglobulin heavy chain; processed gene.  
SOURCE Mouse spleen hybridoma (haplotype 1gh) cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 357)  
AUTHORS Kofler,R.  
TITLE A new murine Ig VH gene family  
JOURNAL J. Immunol. 140 (11), 4031-4034 (1988)  
MEDLINE 88229103  
PUBMED 3131427  
FEATURES  
SOURCE location/Qualifiers  
1..357  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
1..>357  
/note="Ig heavy chain MRL-DNA4 V-region precursor"  
/codon\_start=1  
/protein\_id="AAA38634.1"  
/db\_xref="GI:196221"  
/translation="MLGLKWFVFVFCYGVCHVCLVETGGGLVQPKSLKLSKPAS  
GFSNTNANMVRQAPKGLKLEWVARIRSKSNNTYVADSVKDRFTISRDSOSMLYL  
QNNLKTEDTAMYYCVR"  
1..57  
/note="Ig heavy chain V-region signal peptide"  
mat\_peptide 58..>357  
sig\_peptide  
BASE COUNT 99 a 65 c 92 g 101 t  
ORIGIN  
Query Match 39.0%; Score 144; DB 10; Length 357;  
Best Local Similarity 99.5%; Pred. No. 1.5e-76;  
Matches 194; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 106 TGGGTCCGCCAGGCTCCAGGAAGGTTTGAATGGTGTCTGCATTAAGAATAAAGT 165  
|||||  
Db 163 TGGGTCCGCCAGGCTCCAGGAAGGTTTGAATGGTGTCTGCATTAAGAATAAAGT 222  
|||||

QY 166 AATAATTATGCAACATTTTATGCCGATTCAGTGAAGACAGGTTACCATCTCCAGAGAT 225  
|||||  
Db 223 AATAATTATGCAACATTTTATGCCGATTCAGTGAAGACAGGTTACCATCTCCAGAGAT 282  
|||||

QY 226 GATTACAAACGATGCTCTATCTGCAATGAACAACCTTGAAGAGACACAGCCATG 285  
|||||  
Db 283 GATTACAAACGATGCTCTATCTGCAATGAACAACCTTGAAGAGACACAGCCATG 342  
|||||

QY 286 TATTACTGTGTGAGA 300  
|||||  
Db 343 TATTACTGTGTGAGA 357  
|||||

RESULT 8  
LOCUS MUSIHCVRA 360 bp mRNA linear ROD 26-MAR-2002  
DEFINITION Mouse mRNA for immunoglobulin heavy chain variable region, partial  
sequence.  
ACCESSION D50376.1 GI:804892  
VERSION D50376.1  
KEYWORDS immunoglobulin heavy chain variable region.  
SOURCE Mus musculus (strain BALB/c) hybridoma producing anti-human  
procollagenase antibody, cell-line K1E6 cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
1	(bases 1 to 474)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognatha; Muridae; Murinae; Mus.				
2	Kofler, R., Strohal, R., Balders, R.S., Johnson, M.E., Noonan, D.J., Duchosal, M.A., Dixon, F.J., and Theofilopoulos, A.N.	Immunoglobulin kappa 1 light chain variable region gene complex organization and immunoglobulin genes encoding anti-DNA autoantibodies in lupus mice	J. Clin. Invest. 82 (3), 852-860 (1988)	88331394	3138286	
3	Draft entry and computer-readable sequence [1] kindly submitted by R. Kofler		28-Jul-1988.			
FEATURES						
Source						
1.	474	Location/Qualifiers				
/organism="Mus musculus"						
/db_xref="taxon:10090"						
1.	>474					
/product="Ilg2ak mRNA"						
52.	>474					
/note="Ilg2ak heavy chain precursor"						
/codon_start=1						
/protein_id="AA38849.1"						
/db_xref="GI:196952"						
/translation="MLGLKMFVFPVYQGVHCHVQLVETGGGLVQKGSILKSPAS						
GFSTNTAKMMVROAPGKGLVAVIRKSSNNATVYADSKDRFTISKDSDSMLYL						
QNNLKTEDTAMVYCVBDANMSAMFAYWGQTLVTVA"						
mat_peptide						
52.	>474	/product="Ilg heavy chain"				
52.	111					
/note="Ilg heavy chain signal peptide"						
BASE COUNT	120 a	95 c	119 g	140 t		
ORIGIN	Chromosome 12.					
Query Match						
Best Local Similarity	39.0%;	Score 144;	DB 10;	Length 474;		
Matches 194;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;		
QY	106 TGGGTCGCCGACGGCTCCAGAAAGGTTTGGATGGGTGCTCCGATTAAGATAAAAGT	165				
Db	214 TGGGTCGCCGACGGCTCCAGAAAGGTTTGGATGGGTGCTCCGATTAAGATAAAAGT	273				
QY	166 AATAATTATGCAACATTTTATATGCGATTCACTGGAAGACAGGTTCACCATCTCCAGACAT	225				
Db	274 AATAATTATGCAACATTTTATATGCGATTCACTGGAAGACAGGTTCACCATCTCCAGACAT	333				
QY	226 GATTTCACAACGATGCTATCTGCAATGAAACAATTGAAAACGTGAAGACACAGCCATG	285				
Db	334 GATTTCACAACGATGCTATCTGCAATGAAACAATTGAAAACGTGAAGACACAGCCATG	393				
QY	286 TATTACTGTGTGAGA	300				
Db	394 TATTACTGTGTGAGA	408				
RESULT 10						
LOCUS	AF064446	661 bp	DNA	linear	ROD 04-JAN-2000	
DEFINITION	Mus musculus immunoglobulin heavy chain variable region (VH10.3)					
ACCESSION	AF064446					
VERSION	AF064446.1					
KEYWORDS	GI:3420274					
SOURCE						
ORGANISM	Mus musculus.					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognatha; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 661)					
TITLE	Whitcomb, E.A., Haines, B.B., Parmelee, A.P., Pearlman, A.M. and Broderick, P.H.					
JOURNAL	Germine structure and differential utilization of Igna and Ighb					
J. Immunol.	162 (3), 1541-1550 (1999)					

MEDLINE 99138834  
 PUBMED 9973411  
 REFERENCE 2 (bases 1 to 661)  
 AUTHORS Whitcomb,E.A., Haines,B.B., Parmelee,A.P., Perlman,A.M. and Brodeur,P.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-MAY-1998) Pathology, Tufts University, 136 Harrison Ave, Boston, MA 02111, USA  
 FEATURES  
 source location/Qualifiers  
 1..661  
 /organism="Mus musculus"  
 /strain="BALB/cByJ"  
 /db\_xref="taxon:10090"  
 /chromosome="12"  
 /tissue\_type="liver"  
 /gemline  
 56..>661  
 /gene="Vh10.3"  
 /allele="Vh10.3a"  
 56..63  
 /gene="Vh10.3"  
 /note="octamer"  
 82..87  
 /gene="Vh10.3"  
 /join(<176..221,310..>622)  
 /product="immunoglobulin heavy chain variable region"  
 /gene="Vh10.3"  
 /join(176..221,310..>622)  
 /gene="Vh10.3"  
 /codon\_start=1  
 /product="immunoglobulin heavy chain variable region"  
 /protein\_id="AAC1898.1"  
 /db\_xref="GI:3420275"  
 /translation="MLGLKMFVFFVYQGVCEVQVLETGGGLVQPKGLKLSKAS GFENFTAAHVRQAPGKGLEWVARIRSKSNVATYVADSVKDRFTISRDSQSMVLV OMNLIKTEPDAFYCVR"  
 sig\_peptide 176..220  
 /gene="Vh10.3"  
 <176..221  
 /gene="Vh10.3"  
 /note="leader exon"  
 /join(221,310..>622)  
 /gene="Vh10.3"  
 310..622  
 /gene="Vh10.3"  
 /note="variable region exon"  
 623..661  
 /gene="Vh10.3"  
 /note="recombination signal sequence RSS"  
 BASE COUNT 186 a 129 c 159 g 187 t  
 ORIGIN  
 Query Match 39.0%; Score 144; DB 10; Length 661;  
 Best Local Similarity 99.5%; Pred. No. 1.5e-76;  
 Matches 194; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11  
 MMIGH1 300 bp mRNA linear ROD 05-DEC-1998  
 LOCUS M.musculus mRNA (L1-5A1) for Igh heavy chain V region.  
 DEFINITION X59104  
 ACCESSION X59104.1 GI:51920  
 VERSION I9 heavy chain; I9 variable region; immunoglobulin; rearranged.  
 KEYWORDS house mouse.  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 300)  
 AUTHORS Kavalier,J.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-APR-1991) J. Kavalier, Wistar Institute, 3601 Spruce Street, Philadelphia PA 19104, USA  
 REFERENCE 2 (bases 1 to 300)  
 AUTHORS Kavalier,J.  
 JOURNAL Unpublished  
 FEATURES  
 source location/Qualifiers  
 1..300  
 /organism="Mus musculus"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /cell\_type="B cell"  
 /tissue\_type="hybridoma"  
 1..300  
 /partial  
 /codon\_start=1  
 /product="Igh heavy chain V region"  
 /protein\_id="CAA1850.1"  
 /db\_xref="GI:3980166"  
 /translation="MYQPKSLKLSKASGFTFNANMVRQAPGKLEWVARIR SKSNVATYVADSVKDRFTISRDSQSMVLVOMNLIKTEPDAFYCVDSFHHGCG"  
 CDS  
 BASE COUNT 91 a 67 c 69 g 73 t  
 ORIGIN  
 Query Match 38.8%; Score 143; DB 10; Length 300;  
 Best Local Similarity 99.5%; Pred. No. 6e-76;  
 Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TITLE Germline structure and differential utilization of Igna and Ighb  
VH10 genes  
JOURNAL J. Immunol. 162 (3), 1541-1550 (1999)  
MEDLINE 99138834  
PUBMED 9973411  
REFERENCE 2 (bases 1 to 663)  
AUTHORS Whitcomb,E.A., Haines,B.B., Parmelee,A.P., Perlman,A.M. and Brodeur,P.H.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAY-1998) Pathology, Tufts University, 136 Harrison Ave, Boston, MA 02111, USA  
FEATURES  
source 1..663  
/organism="Mus musculus"  
/strain="BALB/cByJ"  
/db\_xref="taxon:10090"  
/chromosome="12"  
/tissue\_type="liver"  
/germline 57..>663  
/gene="VH10.2"  
/allele="VH10.2a"  
57..64  
/gene="VH10.2"  
/note="octamer"  
83..88  
/gene="VH10.2"  
/join(<178..223,312..>624)  
/gene="VH10.2"  
/product="Immunoglobulin heavy chain variable region"  
/join(178..223,312..>624)  
/gene="VH10.2"  
/codon\_start=1  
/product="Immunoglobulin heavy chain variable region"  
/protein\_id="AAC31897.1"  
/db\_xref="GI:3420273"  
/translation="MLGLKRVFFVYGVGVCEVYGVSSGGLVOPKSLKSCAAS  
GTFNTYANNVYRQAPGKLEVARIRSRNSNNYATYVDSVADRFTISDSDSMYLY  
OMNKLKTEPTARYCVS"  
178..222  
/gene="VH10.2"  
<178..223  
/gene="VH10.2"  
/note="leader exon"  
/join(223,312..>624)  
/gene="VH10.2"  
312..624  
/gene="VH10.2"  
/note="variable region exon"  
625..663  
/gene="VH10.2"  
/note="recombination signal sequence RSS"  
BASE COUNT 182 a 129 c 159 g 193 t  
ORIGIN  
Query Match 38.8%; Score 143; DB 10; Length 663;  
Best Local Similarity 99.5%; Pred. No. 6,1e-76;  
Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 106 TGGGTCGGCCAGGCTCCAGGAAGGGTTTGGAAATGGCTTCGCATTAAGAGTAAAGT 165  
|||||  
DB 428 TGGGTCGGCCAGGCTCCAGGAAGGGTTTGGAAATGGCTTCGCATTAAGAGTAAAGT 487  
QY 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGAT 225  
|||||  
DB 488 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGAT 547  
QY 226 GATTCAAAAGCATGCTCTATCTGCAAAATGACAACCTTGAAGCTGAGACACAGCATG 285  
|||||  
DB 548 GATTCAAAAGCATGCTCTATCTGCAAAATGACAACCTTGAAGCTGAGACACAGCATG 607  
QY 286 TATTACTGTGTGAG 299  
|||||

DB 608 TATTACTGTGTGAG 621  
RESULT 13  
AF072804 299 bp mRNA linear ROD 13-JUL-1998  
LOCUS AF072804  
DEFINITION Mus musculus B.143 anti-DNA immunoglobulin Igm heavy chain variable region mRNA, partial cds.  
ACCESSION AF072804  
VERSION AF072804.1 GI:3309238  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
AUTHORS 1 (bases 1 to 299)  
Krishnan,M.R. and Marlon,T.N.  
TITLE Comparison of the frequencies of arginines in heavy-chain CDR3 of antibodies expressed in the primary B cell repertoires of autoimmune-prone and normal mice  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 299)  
AUTHORS Krishnan,M.R. and Marlon,T.N.  
TITLE Direct Submission  
JOURNAL Submitted (21-JUN-1998) Microbiology and Immunology, The University of Tennessee, Memphis, 858 Madison Ave., Memphis, TN 38163, USA  
FEATURES  
source 1..299  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="B.143"  
/cell\_type="B cell hybridoma generated from LPS-stimulated spleen cells"  
<1..>299  
/codon\_start=3  
/product="anti-DNA immunoglobulin Igm heavy chain variable region"  
/protein\_id="AAC26043.1"  
/db\_xref="GI:3309239"  
/translation="KGSLLSCASGTFNTYANNVYRQAPGKLEVARIRSRNSNNT  
ATYVDSVADRFTISDSDSMYLYOMNKLKTEPTARYCVASYGNFPAVWGQ"  
BASE COUNT 90 a 66 c 67 g 76 t  
ORIGIN  
Query Match 38.2%; Score 141; DB 10; Length 299;  
Best Local Similarity 99.5%; Pred. No. 1e-74;  
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 106 TGGGTCGGCCAGGCTCCAGGAAGGGTTTGGAAATGGCTTCGCATTAAGAGTAAAGT 165  
|||||  
DB 66 TGGGTCGGCCAGGCTCCAGGAAGGGTTTGGAAATGGCTTCGCATTAAGAGTAAAGT 125  
QY 166 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGAT 225  
|||||  
DB 126 AATAATTATGCAACATTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGAT 185  
QY 226 GATTCAAAAGCATGCTCTATCTGCAAAATGACAACCTTGAAGCTGAGACACAGCATG 285  
|||||  
DB 186 GATTCAAAAGCATGCTCTATCTGCAAAATGACAACCTTGAAGCTGAGACACAGCATG 245  
QY 286 TATTACTGTGTG 297  
|||||  
DB 246 TATTACTGTGTG 257  
RESULT 14  
MDIGMVAQ 360 bp mRNA linear ROD 01-APR-1993  
LOCUS MDIGMVAQ  
DEFINITION M.domesticus Igm variable region.  
ACCESSION 222080  
VERSION 222080.1 GI:288665  
KEYWORDS  
anti-DNA antibody; Igm gene; Igm variable region; Immunoglobulin.

SOURCE	Mus musculus domesticus
ORGANISM	Mus musculus domesticus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 360)
TITLE	Tillman,D.M., Jou,N., Hill,R.J. and Marion,T.N.
JOURNAL	Both Igm and Igg anti-Igm antibody are the products of clonally selective B cell stimulation in (NZB x NZW)F1 mice
REFERENCE	J. Exp. Med. 176 (1992) In press
AUTHORS	2 (bases 1 to 360)
TITLE	Direct Submission
JOURNAL	Submitted (23-MAR-1993) Tony N. Marion, Microbiology and Immunology, University of Tennessee, Memphis, 858 Madison Avenue, Memphis, TN, 38163, USA
FEATURES	Location/Qualifiers
source	1..360
gene	/organism="Mus musculus domesticus"
CDS	/strain="(NZB x NZW)F1"
	/isolate="mouse #17"
	/db_xref="taxon:10092"
	/chromosome="12"
	/sex="Female"
	/cell_line="17p.79"
	/cell_type="hybridoma"
	/tissue_type="spleen"
	/dev_stage="somatic variant"
	1..360
	/gene="Igm"
	<1..>360
	/gene="Igm"
	/function="heavy chain variable region for Igm anti-DNA antibody; Vh558 family"
	/note="17p.79 is clonally related to the Igm hybridoma 17p.80 and the Igg hybridoma 17s.161 in the 17ps-c7 clone"
	/codon_start=1
	/product="immunoglobulin variable region"
	/protein_id="CAA80054.1"
	/db_xref="GI:288686"
	/translation="EVQLVADGGGLVOPKGS LKLSCAASGFSFNTNMMNVRAQPKGK LEWVARIRKSNINATYADSVKDRFTISRDDSGSMLEQLNNIKTETDTAYVCMRDD VYAMFAWVGQETLVTSVA"
BASE COUNT	101 a 74 c 93 g 92 t
ORIGIN	
Query Match	37.4%; Score 138; DB 10; Length 360;
Best Local Similarity	99.5%; Pred. No. 7,1e-73;
Matches 188; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	106 TGGGTCGCCGACGCTCCACAGAAAGGGTTTGGAGTGGTTGCTCGCATGAAGATAAAGT 165
DB	106 TGGGTCGCCGACGCTCCACAGAAAGGGTTTGGAGTGGTTGCTCGCATGAAGATAAAGT 165
OY	166 AATATTATGCAACATTATTATGCGCATCCAGTGAAGACAGGTTCACATCTCCAGAGAT 225
DB	166 AATATTATGCAACATTATTATGCGCATCCAGTGAAGACAGGTTCACATCTCCAGAGAT 225
OY	226 GATTTCACAAGCATCTGATTCGCAAAATGAACAACCTGAAAACGTGAGACACAGCCATG 285
DB	226 GATTTCACAAGCATCTGATTCGCAAAATGAACAACCTGAAAACGTGAGACACAGCCATG 285
OY	286 TATTACTGT 294
DB	286 TATTACTGT 294
RESULT 15	
LOCUS	AF023222 242 bp mRNA linear ROD 30-OCT-1997
DEFINITION	Mus musculus clone 185b, family VH10 anti-fluorescein monoclonal
ACCESSION	AF023222 Igm heavy chain mRNA, partial cds.

KEYWORDS	AF023222.1	GI:2570570
SOURCE	Mus musculus.	
ORGANISM	Mus musculus. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 242) Van der Keyl,H., Hsu,C., Tolat,A., Kansil,S., Dalesandro,M.R., Dorer,D.R., Caton,A. and Owen,J.A. Gene family use and somatic mutation in primary and secondary fluorescein-specific igm antibody responses Immunol. Cell Biol. 74 (3), 245-254 (1996)	
TITLE		
JOURNAL	Immunol. Cell Biol. 74 (3), 245-254 (1996)	
MEDLINE	96392936	
PUBMED	8799724	
REFERENCE	2 (bases 1 to 242) van der Keyl,H. and Owen,J. Direct Submission	
AUTHORS	Submitted (08-SEP-1997) BIOLOGY, Haverford College, 370 Lancaster Ave., Haverford, PA 19041, USA	
TITLE		
JOURNAL		
FEATURES	Location/Qualifiers 1..242 /organism="Mus musculus" /strain="BALB/c" /db_xref="taxon:10090" /chromosome="12" /clone="185b, family VH10" <1..>242 /codon_start=1 /product="anti-fluorescein monoclonal igm heavy chain" /protein_id="AA82170.1" /db_xref="GI:2570571" /translation="APSKGLEFVARIRSKSNNTVYADVSKDRFTISRDSQSMLYL QNNLKTEDTAMTYCVKRGVGTGFAWGGSTLVTS"	
BASE COUNT	73 a 45 c 63 g 61 t	
ORIGIN		
Query Match	35.8%; Score 132; DB 10; Length 242;	
Best Local Similarity	99.5%; Pred. No. 3.4e-69;	
Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY 118 GCTCCAGAGAAAGGTTTGGATGGGTCTCGCATAGACGTAAAGTAATTAATTATGCA 177       Db 1 GCTCCAGAGAAAGGTTTGGATGGGTCTCGCATAGAGTAAGTAATTAATTATGCA 60		
QY 178 ACATTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCAAAAGC 237       Db 61 ACATTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCAAAAGC 120		
QY 238 ATGCTCTATCTGCAATAGCAACTTGGAAGAACTGAGACACAGCATCTATTACTGTGTG 297       Db 121 ATGCTCTATCTGCAATAGCAACTTGGAAGAACTGAGACACAGCATCTATTACTGTGTG 180		
QY 298 AGA 300      Db 181 AGA 183		

Search completed: November 27, 2002, 06:25:15  
Job time : 1473.52 secs



Gencore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:54:06 ; Search time 198.197 Seconds  
(without alignments)  
4192.745 Million cell updates/sec

Title: US-09-893-615-86

Perfect score: 369  
Sequence: 1 GAAGTGCATGCTGCTGAGTC.....CCTCAGTCACCGCTCTCTCA 369

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq.101002.\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	132	35.8	330	18	AA158261
2	130	35.2	345	2	Lead binding MAB 6
3	110	29.8	1079	21	AA251115
4	96	26.0	96	20	AA158267
5	90	24.4	90	20	AA158267
6	80	21.7	357	17	AA158267
7	80	21.7	1797	17	AA158267
8	76	20.6	359	21	AA158267
9	69	18.7	69	20	AA158267

10	58	15.7	443	22	AA13178
11	57	15.4	57	20	AA158267
12	57	15.4	57	20	AA158267
13	41	11.1	857	10	AA158267
14	40	10.8	357	22	AA13179
15	38	10.3	375	21	AA140202
16	36	9.8	351	20	AA131382
17	36	9.8	357	17	AA139561
18	36	9.8	440	19	AA122074
19	36	9.8	478	20	AA188433
20	36	9.8	738	15	AA170612
21	36	9.8	1528	5	AA140024
22	36	9.8	1528	5	AA140025
23	36	9.8	1553	16	AA179930
24	35	9.5	42	13	AA123444
25	35	9.5	42	14	AA144176
26	35	9.5	42	17	AA173234
27	35	9.5	42	18	AA172531
28	35	9.5	42	18	AA170515
29	35	9.5	42	18	AA173465
30	35	9.5	42	19	AA138148
31	35	9.5	42	20	AA121887
32	35	9.5	42	20	AA106031
33	35	9.5	48	20	AA131384
34	35	9.5	49	20	AA131379
35	35	9.5	51	22	AA122042
36	35	9.5	52	19	AA118537
37	35	9.5	52	19	AA118573
38	35	9.5	52	19	AA103848
39	35	9.5	52	22	AA121995
40	35	9.5	52	22	AA122024
41	35	9.5	57	20	AA131376
42	35	9.5	57	20	AA131388
43	35	9.5	331	18	AA143738
44	35	9.5	346	18	AA143740
45	35	9.5	357	18	AA158262

#### ALIGNMENTS

RESULT 1	AA158261	standard; cDNA: 330 BP.
ID	AA158261	
AC	AA158261	
XX		
DT	22-AUG-1997	(first entry)
XX		
DE	Lead binding MAB 6F5 heavy chain variable region encoding cDNA.	
XX		
KW	Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic;	
KW	pharmaceutical; health care; skin treatment; pesticide; herbicide;	
KW	heavy metal; ss.	
XX		
OS	Mus musculus.	
XX		
FT	key	Location/Qualifiers
FT	CDS	1..330
XX		
XX	MO9639518-A1.	
XX		
PD	12-DEC-1996.	
XX		
PF	05-JUN-1996;	96MO-US09258.
XX		
PR	10-OCT-1995;	95US-0541373.
PR	05-JUN-1995;	95US-0462798.
XX		
XX	(BION-) BIONEERASKA INC.	
XX		

Murine antibody 1D  
Anti-Staph (HAY) 9  
Anti-Staph (HAY) 9  
Insert coding for  
Humanised murine a  
H. pylori 26 kDa p  
MMS561H M. muscul  
Variable heavy cha  
DNA encoding the h  
EST clone GF196.  
11-6 binding Inhib  
Combined cDNA inse  
mRNA encoding gamm  
Anti-tobacco mosai  
JH specific oligo  
Probe specific for  
Mouse heavy chain  
Probe for mouse he  
Mouse immunoglobul  
Mouse heavy chain  
Probe for the mous  
Probe used to isol  
Mouse heavy chain  
MUSIGHAD mouse Ig  
MUSIGHAD mouse Ig  
Mouse Ig V region  
Mouse Ig V region  
Mouse heavy chain  
Mouse heavy chain  
Mouse heavy chain  
Mouse immunoglobul  
Mouse Ig V region  
MUSIGHAD mouse pr  
MUSIGHAD mouse Ig  
Anti-DNA antibody  
Anti-DNA antibody  
Lead binding MAB 8

PI Lopez O, Murray PJ, Wylie DE;  
XX  
XX WPI: 1997-043140/04.  
DR P-PSDB: AAM01587.  
XX  
XX DNA encoding heavy metal binding polypeptide sequences - used for  
PT detecting, removing, adding or neutralising heavy metals, such as  
PT lead cations  
XX  
PS Claim 1; Page 76; 125pp; English.  
XX  
XX The present sequence encodes the heavy chain variable region for  
CC monoclonal antibody (Mab) 6F5, which immunoreacts with a lead cation.  
CC The sequence was derived from RNA isolated from hybridoma cells from  
CC mouse spleen cells. The protein can be used for binding heavy metals,  
CC such as lead cations. It can be used for detecting, removing, adding  
CC or neutralising the heavy metals in biological and inanimate systems.  
CC It can be used in e.g. aqueous liquid systems, in biological or  
CC environmental systems or in such compositions as perfumes, cosmetics,  
CC pharmaceuticals, health care products, skin treatment products,  
CC pesticides, herbicides, solvents used in the production of semi-  
CC conductor and integrated circuit components and production materials  
CC for electronic components.  
XX  
SQ Sequence 330 BP; 93 A; 73 C; 82 G; 82 T; 0 other;  
Query Match 35.8%; Score 132; DB 18; Length 330;  
Best Local Similarity 99.5%; Pred. No. 2.2e-58;  
Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
QY 121 CCAGAAAGGGTTGGAATGGTTCCTGCATAGAACTAAAGTAATATTATGCAACA 180  
DB 106 CCAGAAAGGGTTGGAATGGTTCCTGCATAGAACTAAAGTAATATTATGCAACA 165  
QY 181 TTTATGCGCATTCAGTGAAGACAGGTTCCACATCTCCAGAGATGATTCACAAAGCATG 240  
DB 166 TATTATGCCGATTCAGTGAAGACAGGTTCCACATCTCCAGAGATGATTCACAAAGCATG 225  
QY 241 CTCTATCTGCAAAATGAACTTGAAGAACTGAGGACACAGCATGTATTACTGTGTGAGA 300  
DB 226 CTCTATCTGCAAAATGAACTTGAAGAACTGAGGACACAGCATGTATTACTGTGTGAGA 285  
QY 301 CGG 303  
DB 286 CGG 288  
XX  
RESULT 2  
AAT58267  
ID AAT58267 standard; cDNA; 345 BP.  
XX  
XX AAT58267;  
XX  
DT 22-AUG-1997 (first entry)  
XX  
DE Lead binding Mab 2B4 heavy chain variable region encoding cDNA.  
XX  
XX Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic;  
KW pharmaceutical; health care; skin treatment; pesticide; herbicide;  
XX heavy metal; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..345  
FT /tag= a  
FT /note= "no stop codon given"  
XX  
XX MO9639518-A1.  
XX  
XX 12-DEC-1996.  
XX  
XX 05-JUN-1996; 96MO-US09258.

XX  
PR 10-OCT-1995; 95US-0541373.  
PR 05-JUN-1995; 95US-0462798.  
XX  
XX (BION-) BIONEERASKA INC.  
XX  
XX Lopez O, Murray PJ, Wylie DE;  
PI  
XX  
XX WPI: 1997-043140/04.  
DR P-PSDB: AAM01593.  
XX  
XX DNA encoding heavy metal binding polypeptide sequences - used for  
PT detecting, removing, adding or neutralising heavy metals, such as  
PT lead cations  
XX  
PS Claim 1; Page 88; 125pp; English.  
XX  
XX The present sequence encodes the heavy chain variable region for  
CC monoclonal antibody (Mab) 2B4, which immunoreacts with a lead cation.  
CC The sequence was derived from RNA isolated from hybridoma cells from  
CC mouse spleen cells. The protein can be used for binding heavy metals,  
CC such as lead cations. It can be used for detecting, removing, adding  
CC or neutralising the heavy metals in biological and inanimate systems.  
CC It can be used in e.g. aqueous liquid systems, in biological or  
CC environmental systems or in such compositions as perfumes, cosmetics,  
CC pharmaceuticals, health care products, skin treatment products,  
CC pesticides, herbicides, solvents used in the production of semi-  
CC conductor and integrated circuit components and production materials  
CC for electronic components.  
XX  
SQ Sequence 345 BP; 100 A; 75 C; 87 G; 83 T; 0 other;  
Query Match 35.2%; Score 130; DB 18; Length 345;  
Best Local Similarity 99.4%; Pred. No. 2.4e-57;  
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
QY 123 AGAAAGGGTTGGAATGGTTCCTGCATAGAAAGTAATATTATGCAACATT 182  
DB 123 AGAAAGGGTTGGAATGGTTCCTGCATAGAAAGTAATATTATGCAACATA 182  
QY 183 TTATGCCGATTCAGTGAAGACAGGTTCCACATCTCCAGAGATGATTCACAAAGCATCT 242  
DB 183 TTATGCCGATTCAGTGAAGACAGGTTCCACATCTCCAGAGATGATTCACAAAGCATCT 242  
QY 243 CTATCTGCAAAATGAACTTGAAGAACTGAGGACACAGCATGTATTACTGTGTGAGACG 302  
DB 243 CTATCTGCAAAATGAACTTGAAGAACTGAGGACACAGCATGTATTACTGTGTGAGACG 302  
QY 303 G 303  
DB 303 G 303  
XX  
RESULT 3  
AAZ51115  
ID AAZ51115 standard; DNA; 1079 BP.  
XX  
XX AAZ51115;  
XX  
DT 05-JUN-2000 (first entry)  
XX  
DE Anti-Her2neu single chain antibody gene.  
XX  
XX Retroviral vector; Spleen Necrosis Virus; SNV;  
KW antibody-envelope fusion protein; retroviral envelope protein;  
KW gene therapy; antigen binding site; single chain antibody; human;  
KW Her2neu cell surface marker; cell specific gene transfer; ds.  
XX  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 14..1015  
FT /tag= a

FT /product- "Anti-Her2neu single chain antibody"  
XX  
PN WO200009730-A2.  
XX  
PD 24-FEB-2000.  
XX  
PF 10-AUG-1999; 99WO-US18141.  
XX  
PR 17-AUG-1998; 98US-0135121.  
XX  
PA (UYJE-) UNIV JEFFERSON THOMAS.  
XX  
PI Dornburg RC;  
XX  
DR WPI: 2000-224358/19.  
DR P-PSDB: AAT70111.  
XX  
PT Cell specific gene transfer using retroviral vectors containing  
XX antibody-envelope fusion proteins and wild type envelope proteins  
XX  
PS Example: Fig 8: 45pp; English.  
XX  
XX The patent discloses a novel retroviral vector, particularly Spleen  
CC Necrosis Virus (SNV) vector, having target cell specificity. The vector  
CC has a targeting envelope which is a chimeric protein consisting of an  
CC antigen binding site of an antibody (e.g. anti-DNP-scfv) or another  
CC peptide that binds to a specific cell surface protein, fused to the  
CC carboxy terminal part of the retroviral envelope protein. The presence  
CC of the wild type envelope protein serves as a helper molecule to  
CC improve or supplement a functional membrane fusion domain. The  
CC antigen binding site replaces the natural viral receptor binding site.  
CC The retroviral vector is used for cell specific gene transfer.  
CC especially in gene therapy. The invention overcomes the restricted host  
CC range limitation of retroviral vectors. The present sequence is a gene  
CC encoding anti-Her2neu single chain antibody (scA). This sequence was used  
CC in the construction of plasmid pAV7 which contains a targeting vector  
CC comprising the anti-Her2neu sca fused to the  
CC SNV-Env(envelope)-TM(transmembrane) coding region.  
XX  
XX Sequence 1079 BP; 303 A; 253 C; 255 G; 268 T; 0 other;  
SO  
Query Match 29.8%; Score 110; DB 21; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 5.7e-47;  
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 189 CGATTAGTGAAGACAGCTTCACATCTCCAGAGATGATTCACAAAGCTGCTATCT 248  
DB 631 CGATTAGTGAAGACAGCTTCACATCTCCAGAGATGATTCACAAAGCTGCTATCT 690  
QY 249 GCAATGACCACTGAAACTGAGACAGCCATCTATTACTGTGTGA 298  
DB 691 GCAATGACCACTGAAACTGAGACAGCCATCTATTACTGTGTGA 740  
RESULT 4  
AAK05582  
ID AAK05582 standard; DNA; 96 BP.  
XX  
AC AAK05582;  
XX  
DT 22-APR-1999 (first entry)  
XX  
DE Anti-Staph (HAY) 96-110 heavy chain variable region encoding DNA.  
XX  
KW Monoclonal antibody; MAb; lipoteichoic acid; gram positive; bacteria;  
KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
XX MAb 96-110; ss.  
XX  
OS Mus sp.  
XX  
FH key Location/Qualifiers  
FT CDS 1..96  
FT FT /\*tag= a

XX  
PN WO9857994-A2.  
XX  
PD 23-DEC-1998.  
XX  
PF 16-JUN-1998; 98WO-US12402.  
XX  
PR 16-JUN-1997; 97US-0049871.  
XX  
PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
XX  
PI Fischer GW, Schuman RF, Stinson JL, Wong H;  
XX  
DR WPI: 1999-095329/08.  
DR P-PSDB: AAM94738.  
XX  
XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT infections caused by gram positive bacteria  
XX  
PS Claim 17; Fig 12; 150pp; English.  
XX  
XX The invention relates to a monoclonal antibody (MAb) to lipoteichoic  
CC acid of gram positive bacteria, where the MAb is a chimeric  
CC immunoglobulin comprising at least part of a human immunoglobulin  
CC constant region and at least part of a non-human immunoglobulin variable  
CC region having specificity to lipoteichoic acid of gram positive bacteria.  
CC The antibodies bind to whole bacteria and enhance phagocytosis and  
CC killing of the bacteria and enhance protection from lethal infection. The  
CC antibodies or peptides (encoded by a DNA of the variable region of  
CC anti-lipoteichoic acid antibody or characterised by amino acids  
CC corresponding to one or more of the complementarily determining regions  
CC (CDRs) of the variable region of the antibody) can be used for treating  
CC or preventing infections caused by gram positive bacteria. They can also  
CC be used for the diagnosis of gram positive bacterial infections.  
CC Sequences AAK05579-83 represent gene fragments encoding the heavy chain  
CC variable regions of the anti-lipoteichoic antibody 96-100.  
XX  
XX Sequence 96 BP; 33 A; 21 C; 19 G; 23 T; 0 other;  
SO  
Query Match 26.0%; Score 96; DB 20; Length 96;  
Best Local Similarity 100.0%; Pred. No. 1.1e-39;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 205 AGGTCACCATCTCCACAGATGATTCACAAAGCATGCTATCTGCAATGACCAACTTG 264  
DB 1 AGGTCACCATCTCCACAGATGATTCACAAAGCATGCTATCTGCAATGACCAACTTG 60  
QY 265 AAAACTGAGGACACAGCCATGTATTACTGTGTGAGA 300  
DB 61 AAAACTGAGGACACAGCCATGTATTACTGTGTGAGA 96  
RESULT 5  
AAK05579  
ID AAK05579 standard; DNA; 90 BP.  
XX  
AC AAK05579;  
XX  
DT 22-APR-1999 (first entry)  
XX  
DE Anti-Staph (HAY) 96-110 heavy chain variable region encoding DNA.  
XX  
KW Monoclonal antibody; MAb; lipoteichoic acid; gram positive; bacteria;  
KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
XX MAb 96-110; ss.  
XX  
OS Mus sp.  
XX  
FH key Location/Qualifiers  
FT CDS 1..90  
FT FT /\*tag= a

```
PN MO9857994-A2.
XX
XX 23-DEC-1998.
XX
XX 16-JUN-1998; 98WO-US12402.
XX
XX 16-JUN-1997; 97US-0049871.
XX
XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX Fischer GW, Schuman RF, Stinson JL, Wong H;
XX WPI: 1999-095329/08.
XX P-PSDB: AAW94735.
XX
XX New antibodies to lipoteichoic acid of gram positive bacteria - used
XX PT to develop products for the diagnosis, prevention and treatment of
XX PT infections caused by gram positive bacteria
XX
XX Claim 17; Fig 12; 150pp; English.
XX
XX The invention relates to a monoclonal antibody (MAb) to lipoteichoic
XX CC acid of gram positive bacteria, where the MAb is a chimeric
XX CC immunoglobulin comprising at least part of a human immunoglobulin
XX CC constant region and at least part of a non-human immunoglobulin variable
XX CC region having specificity to lipoteichoic acid of gram positive bacteria.
XX CC The antibodies bind to whole bacteria and enhance phagocytosis and
XX CC killing of the bacteria and enhance protection from lethal infection. The
XX CC antibodies or peptides (encoded by a DNA of the variable region of
XX CC anti-lipoteichoic acid antibody or characterised by amino acids
XX CC corresponding to one or more of the complementarity determining regions
XX CC (CDRs) of the variable region of the antibody) can be used for treating
XX CC or preventing infections caused by gram positive bacteria. They can also
XX CC be used for the diagnosis of gram positive bacterial infections.
XX CC Sequences AAX05519-83 represent gene fragments encoding the heavy chain
XX CC variable regions of the anti-lipoteichoic antibody 96-100.
XX
XX Sequence 90 BP; 20 A; 17 C; 28 G; 25 T; 0 other;
XX
XX Query Match 24.4%; Score 90; DB 20; Length 90;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-36;
XX Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GAAGTGATGCTGGTGAGCTGTGGAGATTGTGACGCCCTAAAGGCTCATTTGAACCTC 60
XX |
XX DB 1 GAAGTGATGCTGGTGAGCTGTGGAGATTGTGACGCCCTAAAGGCTCATTTGAACCTC 60
XX
XX QY 61 TCATGTGCAGCCTCTGGATTCACTTCAAT 90
XX |
XX DB 61 TCATGTGCAGCCTCTGGATTCACTTCAAT 90
XX
XX RESULT 6
XX AAT15725
XX ID AAT15725 standard; CDNA: 357 BP.
XX
XX AC AAT15725;
XX
XX DT 25-JUN-1996 (first entry)
XX
XX DE 3B1 heavy chain variable region DNA from PCIB4613.
XX
XX KW delta endotoxin; Bacillus thuringiensis; western corn rootworm;
XX KW MCRM; maize; pesticide; brush border membrane vesicle; monoclonal;
XX KW antibody; ss.
XX
XX OS Insecta sp.
XX
XX MO9600783-A1.
XX
XX PN 11-JAN-1996.
XX
XX PD
XX
XX PF 20-JUN-1995; 95WO-IB00497.
```

```
XX
XX 28-JUN-1994; 94US-0267641.
XX
XX (CIBA ) CIBA GEIGY AG.
XX
XX PI Carozzi NB, Koziel MG;
XX
XX WPI: 1996-077494/08.
XX P-PSDB: AAR90829.
XX
XX New monoclonal antibodies which bind insect gut proteins - used
XX PT partic. with toxin moieties for the control of insect pests. partic.
XX PT in plants
XX
XX Claim 8; Page 50-51; 106pp; English.
XX
XX AAT15725-35 are DNA mols. encoding a monoclonal antibody or a binding
XX CC fragment. MAbs were produced by using insect guts, partic. insect
XX CC brush border membranes (BBMs), esp. corn rootworm, as antigen;
XX CC immunising a donor animal with the antigen; isolating immunocompetent B
XX CC cells from the immunised animal; fusing B cells with a tumour cell line;
XX CC isolating the fused cells, culturing them and cloning positive hybrid
XX CC cells; and screening the hybrid cells for prodn. of the required MAbs.
XX CC The MAbs bind to the gut of a target insect but do not bind to mammalian
XX CC BBMs. The DNA sequence can be operably linked to a toxin moiety, esp.
XX CC selected from e.g. Bacillus toxins, pseudomonas exotoxin and phytolectin,
XX CC etc.. The Abs are useful for control of insect pests, e.g. Coleoptera,
XX CC Diptera, Hymenoptera and Lepidoptera. The pesticidal compsn. is pref.
XX CC applied to a plant, e.g. maize.
XX
XX Sequence 357 BP; 102 A; 76 C; 91 G; 88 T; 0 other;
XX
XX Query Match 21.7%; Score 80; DB 17; Length 357;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-31;
XX Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 15 GGAGTCTGTGGAGGATGCTGCAGCCCTAAAGGCTCATTTGAACCTCTCATGTGCAGCCTC 74
XX |
XX DB 15 GGAGTCTGTGGAGGATGCTGCAGCCCTAAAGGCTCATTTGAACCTCTCATGTGCAGCCTC 74
XX
XX QY 75 TGGATTCACTTCAATTAAT 94
XX |
XX DB 75 TGGATTCACTTCAATTAAT 94
XX
XX RESULT 7
XX AAT15733
XX ID AAT15733 standard; CDNA: 1797 BP.
XX
XX AC AAT15733;
XX
XX DT 25-JUN-1996 (first entry)
XX
XX DE 3B1 single chain antibody DNA from PCIB4631.
XX
XX KW delta endotoxin; Bacillus thuringiensis; western corn rootworm;
XX KW MCRM; maize; pesticide; brush border membrane vesicle; monoclonal;
XX KW antibody; ss.
XX
XX OS Insecta sp.
XX
XX MO9600783-A1.
XX
XX PN 11-JAN-1996.
XX
XX PD
XX
XX PF 20-JUN-1995; 95WO-IB00497.
XX
XX PR 28-JUN-1994; 94US-0267641.
XX
XX PA (CIBA ) CIBA GEIGY AG.
XX
XX PI Carozzi NB, Koziel MG;
```

DR WPI: 1996-077494/08.  
 DR P-PSDB: AAR90837.  
 XX  
 PT New monoclonal antibodies which bind insect gut proteins - used  
 PT partic. with toxin moieties for the control of insect pests, partic.  
 PT in plants  
 XX  
 PS Claim 8: Page 68-72; 106pp; English.  
 XX  
 CC AA15725-35 are DNA mols. encoding a monoclonal antibody or a binding  
 CC fragment. Mabs were produced by using insect guts, partic. insect  
 CC brush border membranes (BBMs), esp. corn rootworm, as antigen;  
 CC immunising a donor animal with the antigen; isolating immunocompetent B  
 CC cells from the immunised animal; fusing B cells with a tumour cell line;  
 CC isolating the fused cells, culturing them and cloning positive hybrid  
 CC cells; and screening the hybrid cells for prodn. of the required Mabs.  
 CC The Mabs bind to the gut of a target insect but do not bind to mammalian  
 CC BBMs. The DNA sequence can be operably linked to a toxin moiety, esp.  
 CC selected from e.g. Bacillus toxins, Pseudomonas exotoxin and phytoalexin,  
 CC etc.. The Abs are useful for control of insect pests, e.g. Coleoptera,  
 CC Diptera, Hymenoptera and Lepidoptera. The pesticidal compn. is pref.  
 CC applied to a plant, e.g. maize.  
 XX  
 SQ Sequence 1797 BP; 455 A; 488 C; 452 G; 402 T; 0 other;  
 Query Match 21.7%; Score 80; DB 17; Length 1797;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-31;  
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 15 GGAGTGTGTGAGAGATTGGTGCAGCCCTTAAGGTCATTGCAATCTCATGTGCAGCCTC 74  
 DB 477 GGAGTGTGTGAGAGATTGGTGCAGCCCTTAAGGTCATTGCAATCTCATGTGCAGCCTC 536  
 QY 75 TGGATTCACCTTCATTAACCT 94  
 DB 537 TGGATTCACCTTCATTAACCT 556

RESULT 8  
 AAA44338  
 ID AAA44338 standard; cDNA; 359 BP.  
 AC AAA44338;  
 DT 21-AUG-2000 (first entry)  
 DE Human secreted expressed sequence tag SEQ ID NO:913.  
 XX  
 KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antiproliferative; vulnerrary; antiparkinsonian;  
 KW anticancer; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
 KW tumour; infection; depression; psoriasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200021991-A1.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 15-OCT-1999; 99WO-US24206.  
 XX  
 PR 15-OCT-1998; 98US-0104436.  
 XX

PA (GENE) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
 PI Werberg D, Treacy M, Bowman MR;  
 XX  
 DR WPI: 2000-317938/27.  
 XX  
 PT Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (SESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 XX  
 PS Claim 1: Page 437-438; 803pp; English.  
 XX  
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (SESTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The SESTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytosolic; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antiparkinsonian; antipsoriatic; anticonvulsant; neuroprotective;  
 CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
 CC therapy and in vaccines. The SESTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 359 BP; 91 A; 66 C; 93 G; 109 T; 0 other;  
 Query Match 20.6%; Score 76; DB 21; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-29;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 106 TGGGTCGCGCAGGCTCCAGGAAGGTTTGGATGGGTTCTCCATTAACAGTAAGAAT 165  
 DB 219 TGGGTCGCGCAGGCTCCAGGAAGGTTTGGATGGGTTCTCCATTAACAGTAAGAAT 278  
 QY 166 AATAATTATGCAACAT 181  
 DB 279 AATAATTATGCAACAT 294

RESULT 9  
 AAX05583  
 ID AAX05583 standard; DNA; 69 BP.  
 AC AAX05583;  
 DT 22-APR-1999 (first entry)  
 DE Anti-Staph (HAW) 96-110 heavy chain variable region encoding DNA.  
 XX  
 KW Monoclonal antibody; MAb; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW MAb 96-110; ss.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9857994-A2.  
 XX  
 PH key Location/Qualifiers  
 FT 1..69  
 FT CDS /tag- a  
 XX

```

XX 23-DEC-1998.
XX
XX 16-JUN-1998; 98MO-US12402.
XX
XX 16-JUN-1997; 97US-0049871.
XX
XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
XX Flascher GW, Schuman RF, Stinson JL, Wong H;
XX
XX WPI: 1999-095328/08.
XX
XX P-PSDB: AAW94739.
XX
XX New antibodies to lipoteichoic acid of gram positive bacteria - used
XX to develop products for the diagnosis, prevention and treatment of
XX infections caused by gram positive bacteria
XX
XX Claim 17: Fig 12: 150pp; English.
XX
XX The invention relates to a monoclonal antibody (mAb) to lipoteichoic
XX acid of gram positive bacteria, where the mAb is a chimeric
XX immunoglobulin comprising at least part of a human immunoglobulin
XX constant region and at least part of a non-human immunoglobulin variable
XX region having specificity to lipoteichoic acid of gram positive bacteria.
XX The antibodies bind to whole bacteria and enhance phagocytosis and
XX killing of the bacteria and enhance protection from lethal infection. The
XX antibodies or peptides (encoded by a DNA of the variable region of
XX anti-lipoteichoic acid antibody or characterised by amino acids
XX corresponding to one or more of the complementarity determining regions
XX (CDRs) of the variable region of the antibody) can be used for treating
XX or preventing infections caused by gram positive bacteria. They can also
XX be used for the diagnosis of gram positive bacterial infections.
XX Sequences AAX05579-83 represent gene fragments encoding the heavy chain
XX variable regions of the anti-lipoteichoic antibody 96-100.
XX
XX Sequence 69 BP; 14 A; 19 C; 20 G; 16 T; 0 other;
XX
XX Query Match 18.7%; Score 69; DB 20; Length 69;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-25;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 301 CGGGGGCTTACGAGTACTATGCTATGACACTGCTGGGTCAGAACCTCACTACC 360
XX |||||||
XX 1 CGGGGGCTTACGAGTACTATGCTATGACACTGCTGGGTCAGAACCTCACTACC 60
XX
XX 361 GTCTCTCA 369
XX |||||||
XX 61 GTCTCTCA 69
XX
XX RESULT 10
XX AAD13178
XX ID AAD13178 standard; DNA: 443 BP.
XX
XX AAD13178:
XX
XX 16-OCT-2001 (first entry)
XX
XX Murine antibody 1D9 heavy chain variable region encoding DNA.
XX
XX Murine: humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;
XX neuroprotective; immunosuppressive; human immunodeficiency virus;
XX HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;
XX inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;
XX multiple sclerosis; atherosclerosis; restenosis; asthma;
XX anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;
XX fibrotic disease; angioptosis; acquired immune deficiency syndrome;
XX AIDS; inflammatory glomerulopathy; vascular intervention;
XX neointimal hyperplasia; antibody 1D9 heavy chain variable region;
XX ds.
XX Mus sp.
XX

```

```

XX Key Location/Qualifiers
XX CDS 1..443
XX /*tag= a
XX /product= "Murine antibody 1D9 heavy
XX chain variable region"
XX /transl_except= (pos:442..443, aa:val)
XX /note= "CDS does not include start and stop codon"
XX /partial
XX 1..57
XX /*tag= b
XX mat_peptide 58..443
XX /*tag= c
XX /product= "Murine mature antibody 1D9 kappa
XX light chain variable region"
XX 58..411
XX misc_feature /*tag= d
XX /*note= "Encodes variable region"
XX /*tag= e
XX /*note= "Encodes a part of the constant region"
XX
XX W0200157226-A1.
XX
XX 09-AUG-2001.
XX
XX 02-FEB-2001; 2001MO-US03537.
XX
XX 03-FEB-2000; 2000US-0497625.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;
XX WPI: 2001-48888/53.
XX P-PSDB: AAE07033.
XX
XX Humanized immunoglobulin for treating a CC-chemokine receptor
XX 2-mediated disorder in a patient, comprises a binding specificity for
XX CCR2, and a non-human antigen binding region and human immunoglobulin
XX
XX Claim 24: Fig 21: 183pp; English.
XX
XX The patent discloses a humanised antibody or its antigen-binding
XX fragment, having binding specificity for CC-chemokine receptor 2
XX (CCR2), comprising an antigen binding region of non-human origin
XX and at least a portion of an immunoglobulin of human origin. The
XX humanised antibodies are useful for inhibiting the interaction of
XX a cell expressing CCR2. They are useful for inhibiting or treating
XX HIV infection. The proteins of the invention are useful for inhibiting
XX leukocyte trafficking, for treating CCR2-mediated disorders such as
XX inflammatory disorder, autoimmune disorders such as rheumatoid
XX arthritis and multiple sclerosis, atherosclerosis and atherosclerosis,
XX and for inhibiting restenosis. They are useful in therapy or diagnosis,
XX and in the manufacture of a medicament for treating CCR2 mediated
XX disease. They are also useful for treating allergy, anaphylaxis,
XX malignancy, chronic and acute inflammation, histamine and IgE-
XX mediated allergic reaction, shock, stenosis, allograft rejection,
XX fibrotic disease, asthma, inflammatory glomerulopathies, acquired
XX immune deficiency syndrome (AIDS), restenosis associated with vascular
XX intervention. Including angioptosis and/or stent placement in a mammal.
XX Humanised antibodies are also useful for inhibiting narrowing of the
XX lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of
XX a vessel in a mammal, preferably associated with vascular intervention.
XX The present sequence is a DNA encoding murine antibody 1D9 heavy chain
XX variable region.
XX
XX Sequence 443 BP; 117 A; 97 C; 108 G; 121 T; 0 other;
XX
XX Query Match 15.7%; Score 58; DB 22; Length 443;
XX Best Local Similarity 100.0%; Pred. No. 5.4e-20;
XX Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 25 GGAGATTGTCACGCTAAAGGCTGATTGAACTGTCATGCGACCTTGATTCA 82  
|||||  
Db 82 GGAGATTGTCACGCTAAAGGCTGATTGAACTGTCATGCGACCTTGATTCA 139

## RESULT 11

AA05580 ID AAX05580 standard; DNA: 57 BP.

AC AAX05580;

DT 22-APR-1999 (first entry)

DE Anti-Staph (HAY) 96-110 heavy chain variable region encoding DNA.

KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;

KW Mab 96-110; ss.

OS Mus sp.

FN Key Location/Qualifiers

FT CDS 1..57

FT /\*tag= a

PE 16-JUN-1998; 98WO-US12402.

PR 16-JUN-1997; 97US-0049871.

PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PI Fischer GW, Schuman RF, Stinson JL, Wong H;

DR WPI: 1999-095329/08.

DR P-PSDB: AAW94736.

PT New antibodies to lipoteichoic acid of gram positive bacteria - used  
to develop products for the diagnosis, prevention and treatment of  
infections caused by gram positive bacteria

PS Claim 17; Fig 12; 150pp; English.

CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
acid of gram positive bacteria, where the Mab is a chimeric

CC immunoglobulin comprising at least part of a human immunoglobulin  
constant region and at least part of a non-human immunoglobulin variable

CC region having specificity to lipoteichoic acid of gram positive bacteria.  
The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The  
antibodies or peptides (encoded by a DNA of the variable region of

CC anti-lipoteichoic acid antibody or characterised by amino acids  
corresponding to one or more of the complementarity determining regions

CC (CDRs) of the variable region of the antibody) can be used for treating  
or preventing infections caused by gram positive bacteria. They can also

CC be used for the diagnosis of gram positive bacterial infections.  
Sequences AAX05579-83 represent gene fragments encoding the heavy chain

CC variable regions of the anti-lipoteichoic antibody 96-100.

XX Sequence 57 BP; 13 A; 12 C; 19 G; 13 T; 0 other;

SO Query Match 15.4%; Score 57; DB 20; Length 57;

Best Local Similarity 100.0%; Pred. No. 1.8e-19;

Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 91 AACTACGCCATGAATTGGTCCGACGCTCCAGAAAGGTTTGATGGTTCCT 147  
|||||  
Db 1 AACTACGCCATGAATTGGTCCGACGCTCCAGAAAGGTTTGATGGTTCCT 57

## RESULT 12

AA05581 ID AAX05581 standard; DNA: 57 BP.

AC AAX05581;

DT 22-APR-1999 (first entry)

DE Anti-Staph (HAY) 96-110 heavy chain variable region encoding DNA.

KW Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

KW Immunoglobulin; phagocytosis; infection; epitope; peptide mimic;

KW Mab 96-110; ss.

OS Mus sp.

FN Key Location/Qualifiers

FT CDS 1..57

FT /\*tag= a

PE 16-JUN-1998; 98WO-US12402.

PR 16-JUN-1997; 97US-0049871.

PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

PI Fischer GW, Schuman RF, Stinson JL, Wong H;

DR WPI: 1999-095329/08.

DR P-PSDB: AAW94737.

PT New antibodies to lipoteichoic acid of gram positive bacteria - used  
to develop products for the diagnosis, prevention and treatment of  
infections caused by gram positive bacteria

PS Claim 17; Fig 12; 150pp; English.

CC The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
acid of gram positive bacteria, where the Mab is a chimeric

CC immunoglobulin comprising at least part of a human immunoglobulin  
constant region and at least part of a non-human immunoglobulin variable

CC region having specificity to lipoteichoic acid of gram positive bacteria.  
The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The  
antibodies or peptides (encoded by a DNA of the variable region of

CC anti-lipoteichoic acid antibody or characterised by amino acids  
corresponding to one or more of the complementarity determining regions

CC (CDRs) of the variable region of the antibody) can be used for treating  
or preventing infections caused by gram positive bacteria. They can also

CC be used for the diagnosis of gram positive bacterial infections.  
Sequences AAX05579-83 represent gene fragments encoding the heavy chain

CC variable regions of the anti-lipoteichoic antibody 96-100.

XX Sequence 57 BP; 24 A; 8 C; 10 G; 15 T; 0 other;

SO Query Match 15.4%; Score 57; DB 20; Length 57;

Best Local Similarity 100.0%; Pred. No. 1.8e-19;

Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 148 CGCATGAAGTAAGTAATTAATGCAACATTTATGCCGATTCAGTGAAGAC 204  
|||||

Db 1 CGCATGAAGTAAGTAATTAATGCAACATTTATGCCGATTCAGTGAAGAC 57

## RESULT 13

AA05580 ID AAN90301 standard; DNA: 857 BP.

AC AAN90301;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Insert coding for a heavy chain murine variable region.  
 XX  
 KW Recombinant DNA; chimeric monoclonal antibody; light chain; heavy chain;  
 KW variable region; human carcinoembryonic antigen; cell line CE25.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 322..844  
 FT /\*tag= a  
 FT 370..475  
 FT Intron /\*tag= b  
 XX  
 PN EP323806-A.  
 XX  
 PD 12-JUL-1989.  
 XX  
 PF 28-DEC-1988; 88EP-0810898.  
 XX  
 PR 05-JAN-1988; 88GB-0000077.  
 XX  
 PA (CIBA ) CIBA-GEIGY AG.  
 XX  
 PI Hardman N, Gill LL, de Winter RFJ, Wagner K, Heusser C;  
 XX  
 DR WPI: 1989-200701/28.  
 DR P-PSDB: AAP90480.  
 XX  
 PT Chimeric monoclonal antibody to human carcinoembryonic antigen  
 PT - consisting of variable regions of mouse origin and human constant  
 PT regions, for cancer diagnosis and therapy.  
 XX  
 PS Claim 30; page 37; 53pp; English.  
 XX  
 CC The recombinant DNA is an insert coding for a heavy chain murine  
 CC variable region, which is specific for human carcinoembryonic antigen  
 CC and originates from genomic DNA of the cell line CE 25. See also  
 CC AAP90480.  
 CC  
 SO Sequence 857 BP; 225 A; 186 C; 210 G; 236 T; 0 other;  
 XX  
 Query Match 11.8%; Score 41; DB 10; Length 857;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-11;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 329 TGGACTACTGGGTCAGGAAGCACTCACTCACCGTCTCTCA 369  
 DB 804 TGGACTACTGGGTCAGGAAGCACTCACTCACCGTCTCTCA 844  
 XX  
 RESULT 14  
 AAD13179  
 ID AAD13179 standard; DNA: 357 BP.  
 AC AAD13179;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Humanised murine antibody heavy chain 1D9RHA DNA.  
 XX  
 KW Murine; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
 KW multiple sclerosis; atherosclerosis; restenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angiodysplasia; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention;  
 KW neointimal hyperplasia; antibody ID9 heavy chain; 1D9RHA; ds.  
 XX  
 OS Chimeric - Mus sp.

OS Chimeric - Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..357  
 FT /\*tag= a  
 FT /product= "humanised murine antibody heavy chain  
 FT region, 1D9RHA"  
 FT /note= "CDS does not include start and stop codon"  
 FT /partial  
 XX  
 PN W0200157226-A1.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 02-FEB-2001; 2001WO-US03537.  
 XX  
 PR 03-FEB-2000; 2000US-0497625.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 XX  
 DR WPI: 2001-488888/53.  
 DR P-PSDB: AAE07034.  
 XX  
 PT Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 PT CCR2, and a non-human antigen binding region and human immunoglobulin  
 PT -  
 XX  
 PS Claim 64; Fig 23; 183pp; English.  
 XX  
 CC The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherosclerosis and arteriosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IGE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angiodysplasia and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is a DNA encoding humanised murine antibody heavy  
 CC chain region, 1D9RHA. 1D9RHA sequence consists of the complementary  
 CC determining regions (CDRs) of the murine ID9 antibody heavy chain  
 CC variable (VH) region genetically inserted into the framework regions  
 CC (FRs) of the human 484/CL antibody VH region.  
 XX  
 SO Sequence 357 BP; 99 A; 79 C; 95 G; 84 T; 0 other;  
 XX  
 Query Match 10.8%; Score 40; DB 22; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 106 TGGGTCCGCGAGCTCCAGGAAGGTTTGAATGGCTTG 145  
 DB 106 TGGGTCCGCGAGCTCCAGGAAGGTTTGAATGGCTTG 145  
 XX  
 RESULT 15  
 AAA40202  
 ID AAA40202 standard; DNA: 375 BP.



XX	AA040202;
XX	
DT	01-NOV-2000 (first entry)
XX	
XX	
DE	H. pylori 26 kDa protein-binding antibody heavy chain encoding DNA.
XX	
KW	Acid-resistant microorganism; detection; faecal; intestine; infection;
KW	monoclonal antibody; heavy chain; ds.
OS	Unidentified.
PN	W0200026671-A1.
PD	
PD	11-MAY-2000.
PF	
PF	29-OCT-1999; 99WO-EP08212.
PR	
PR	29-OCT-1998; 98EP-0120517.
XX	06-NOV-1998; 98EP-0120687.
XX	
PA	(CONN-) CONNEX GMBH.
PI	
PI	Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;
PI	Ringels A;
DR	
DR	WPI: 2000-365747/31.
P-PSDB:	AMBI0022.
PT	
PT	Detecting infection by acid-fast microbes for diagnosis of Helicobacter
PT	pylori, comprises reacting a faecal sample with two binding reagents for
PT	antigens that survive intestinal passage
XX	
XX	
PS	Disclosure: Flg 4; 84pp: German.
CC	
CC	This invention describes a novel method for the detection of a mammalian
CC	infection by an acid-resistant microorganism (A) by treating a faecal
CC	sample with at least two different monoclonal antibodies (Mab) (or their
CC	fragments or derivatives) or aptamers (collectively (I)) and detecting
CC	formation of a complex (C) between (I) and the corresponding antigen of
CC	(A). The first and second (I) bind to epitopes of different antigens
CC	(Ae). These epitopes are present, after passage through the intestines,
CC	in at least some mammals, and have either: (i) their native structure;
CC	or (ii) a structure against which an antibody is produced by an animal
CC	infected or immunized with (A), or its extract, lysate, derived protein
CC	or fragment, or with a synthetic peptide. Practically all mammals display
CC	at least one of the specified epitopes. The method is used to detect
CC	infection by acid-fast bacteria, particularly of the genera Helicobacter,
CC	Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,
CC	M. tuberculosis, C. jejuni and C. pylori. (I) may also be used
CC	therapeutically. The method is direct and non-invasive, and provides an
CC	expensive and easily standardizable diagnosis, despite possible
CC	degradation of antigens during passage through the intestines. This
CC	sequence encodes the H. pylori 26 kDa protein-binding antibody
CC	(DMS ACC2255) heavy chain which is used to illustrate the method of the
CC	invention.
XX	
XX	Sequence 375 BP; 102 A; 86 C; 89 G; 98 T; 0 other;
Query Match	10.3%; Score 38; DB 21; Length 375;
Best Local Similarity	100.0%; Pred. No. 1.3e-09;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
263 TGAATACTGAGACACAGCCATGTATTACTGTGTAGA 300	
266 TGAATACTGAGACACAGCCATGTATTACTGTGTAGA 303	

Search completed: November 27, 2002, 05:37:40  
Job time : 207.697 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 05:31:20 ; Search time 1527.56 Seconds  
(without alignments)  
3912.201 Million cell updates/sec

Title: US-09-893-615-86

Perfect score: 369  
Sequence: 1 GAAGTGCATGCTGCTGAGTCA.....CCTCAGTCACTGCTGCTCA 369

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: gb\_gsa:\*  
18: em\_gsa\_hum:\*  
19: em\_gsa\_inv:\*  
20: em\_gsa\_pln:\*  
21: em\_gsa\_vrt:\*  
22: em\_gsa\_fun:\*  
23: em\_gsa\_mam:\*  
24: em\_gsa\_mus:\*  
25: em\_gsa\_other:\*  
26: em\_gsa\_pro:\*  
27: em\_gsa\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	33.1	865	13	BI150936 602916745
2	88	23.8	932	12	BF135929 601781259
3	61	16.5	359	10	BB843560 BB843560
4	39	10.6	365	10	AW403942 UI-HF-BK0
5	36	9.8	565	12	BG686641 602637877
6	35	9.5	552	17	BH319627 CH230-119

7	35	9.5	666	13	BG963642
8	35	9.5	781	13	BG967007
9	35	9.5	861	13	BI455668
10	35	9.5	862	13	BI102617
11	35	9.5	864	10	BE309336
12	35	9.5	1010	14	BO921928
13	34	9.2	567	14	BQ266826
14	34	9.2	837	13	BG966355
15	34	9.2	963	12	BE578938
16	33	8.9	560	12	BF581663
17	33	8.9	1384	11	AK002875
18	32	8.7	348	10	AW402831
19	32	8.7	361	12	BF844111
20	32	8.7	604	10	BE367979
21	32	8.7	616	12	BF136295
22	32	8.7	639	10	BE371136
23	32	8.7	802	12	BF137216
24	32	8.7	819	13	BG966397
25	32	8.7	839	14	BO952276
26	32	8.7	883	13	BI151077
27	32	8.7	941	12	BF138189
28	31	8.4	676	10	BE369087
29	31	8.4	840	13	BI690298
30	31	8.4	890	12	BG868002
31	31	8.4	987	12	BF577496
32	30	8.1	425	12	BE578521
33	30	8.1	724	12	BF168514
34	30	8.1	851	12	BF663091
35	30	8.1	911	10	BE286958
36	29	7.9	332	12	BF155433
37	29	7.9	650	12	BF579001
38	29	7.9	663	10	BB586011
39	29	7.9	690	13	BG968682
40	29	7.9	718	12	BF136279
41	29	7.9	892	12	BF976010
42	29	7.9	927	12	BF179165
43	29	7.9	928	12	BF531379
44	29	7.9	955	14	BQ947728
45	29	7.9	961	14	BQ711482

#### ALIGNMENTS

RESULT 1  
BI150936  
LOCUS 865 bp mRNA linear EST 05-JUL-2001  
DEFINITION 602916745F1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:5067346 5',  
ACCESSION BI150936  
VERSION BI150936.1 GI:14610937  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgaps-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LNAK1181 row: 1 column: 11  
High quality sequence stop: 797.  
Location/Qualifiers 1..865

FEATURES  
source

/organism="Mus musculus"  
/strain="CZECH II"  
/db.xref="taxon:10090"  
/clone="IMAGE:5067346"  
/clone\_lib="NCI\_CGAP\_Lu29"  
/tissue\_type="spontaneous tumor, metastatic to mammary.  
stem cell origin."  
/lab\_host="DH10B"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site:1: SalI;  
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH

BASE COUNT 213 a 229 c 209 g 214 t

ORIGIN

Query Match 33.1%; Score 122; DB 13; Length 865;  
Best Local Similarity 99.4%; Pred. No. 1.8e-56;  
Matches 172; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 106 TGGTCCGCCAGCTCCAGAAAGGGTTGGAAATGGTTGCTGCAGTAAGAAAGT 165  
|||||  
DB 215 TGGTCCGCCAGCTCCAGAAAGGGTTGGAAATGGTTGCTGCAGTAAGAAAGT 274  
|||||

OY 166 AATATTATGACATTTATGCCATTCAGTAAGACAGGTTCCACATCTCCAGAGT 225  
|||||  
DB 275 AATATTATGACATTTATGCCATTCAGTAAGACAGGTTCCACATCTCCAGAGT 334  
|||||

OY 226 GATTCCAAAGCAGTCTATCTGCAATGAACAACATTGAAACAGAGACAC 278  
|||||  
DB 335 GATTCCAAAGCAGTCTATCTGCAATGAACAACATTGAAACAGAGACAC 387  
|||||

RESULT 2  
BF135929 932 bp mRNA linear EST 24-OCT-2000  
LOCUS BF135929  
DEFINITION 601781259p1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4009334 5',  
mRNA sequence.  
ACCESSION BF135929  
VERSION BF135929.1 GI:10974969  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 932)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMNL at:  
<http://image.llnl.gov>  
Plate: LHAM9245 row: e column: 15  
High quality sequence stop: 652.  
Location/Qualifiers  
1. .932  
/organism="Mus musculus"  
/strain="CZECH II"  
/db.xref="taxon:10090"  
/clone="IMAGE:4009334"  
/clone\_lib="NCI\_CGAP\_Lu30"  
/tissue\_type="tumor, metastatic to mammary"  
/lab\_host="DH10B"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site:1: NotI;  
Site:2: SalI; transgenic model MN-1 expression driven by  
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo  
dt. Library constructed by Life Technologies.  
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 257 a 244 c 249 g 182 t

ORIGIN

Query Match 23.8%; Score 88; DB 12; Length 932;  
Best Local Similarity 100.0%; Pred. No. 1.5e-37;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 CATCTCCAGAGATGATTCACAAAGCATGCTATCTGCAAAATGACAACTGAAACTGA 272  
|||||  
DB 166 CATCTCCAGAGATGATTCACAAAGCATGCTATCTGCAAAATGACAACTGAAACTGA 225  
|||||

OY 273 GCACACAGCCATGTTACTGTGTGAGA 300  
|||||  
DB 226 GCACACAGCCATGTTACTGTGTGAGA 253  
|||||

RESULT 3  
BB843560 359 bp mRNA linear EST 26-NOV-2001  
LOCUS BB843560  
DEFINITION BB843560 RIKEN full-length enriched, 6 days neonate spleen Mus  
musculus cDNA clone F430102008 5', mRNA sequence.  
ACCESSION BB843560  
VERSION BB843560.1 GI:17081927  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 359)  
Akimura,T., Arahawa,T., Carninci,P., Furumori,M., Hanagaki,T., Ishii  
Hayatsu,N., Hiramoto,K., Hiroaka,T., Hirozane,T., Imotani,K., Ishii  
Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T.,  
Nakamura,M., Nishii,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,  
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,  
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa  
A., Takahashi,T., Takaku-Akaiira,S., Tanaka,T., Tomaru,A., Toya,T.,  
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: [genome-res@sc.riken.go.jp](mailto:genome-res@sc.riken.go.jp),  
URL:<http://genome-gsc.riken.go.jp/>  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh  
M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi,K., Fujiwara,S., Inoue,K., Toyawa,Y., Iwawa,M., Ohara,E.,  
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura  
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara  
Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome-gsc.riken.go.jp>) for  
further details.  
e mouse tissues.  
Location/Qualifiers  
1. .359  
/organism="Mus musculus"  
/strain="C57BL/6J"

```

/db_xref="taxon:10090"
/clone="F430102008"
/clone_lib="RIKEN full-length enriched, 6 days neonate spleen"
/tissue_type="spleen"
/dev_stage="6 days neonate"
/notes="Vector: pSPOR1, Site_1: SalI; Site_2: NotI. This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
Mol Genet 7: 1967-1978."
BASE COUNT      90 a      69 c      91 g      109 t
ORIGIN

```

```

Query Match      16.5%; Score 61; DB 10; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 106 TGGTCCCGCCAGCTCCAGGAAGGTTGGATGGCTGCTGCCTAAGAACTAAACT 165
      |||||||
DB 229 TGGTCCCGCCAGCTCCAGGAAGGTTGGATGGCTGCTGCCTAAGAACTAAACT 288
QY 166 A 166
      |
DB 289 A 289

```

```

RESULT 4      365 bp      mRNA      linear      EST 16-FEB-2000
LOCUS      AM403942
DEFINITION  UT-HF-BK0-abn-h-04-0-UT.r1 NIH_MGC_36 Homo sapiens cDNA clone
IMAGE:3056887 5', mRNA sequence.
ACCESSION  AM403942
VERSION    AM403942.1 GI:69222910
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens

```

```

REFERENCE 1 (bases 1 to 365)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Eco RI site shown at the beginning of the sequence.
          Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
          cDNA Library Preparation: M.B. Soares Lab
          cDNA Library Arrayed by: M.B. Soares Lab
          DNA Sequencing by: M.B. Soares Lab
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          www.bio.llnl.gov/dbp/image/image.html
          Seq primer: M13 forward

```

```

FEATURES
source
location/Qualifiers
1..365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3056887"

```

```

/clone_lib="NIH_MGC_36"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/notes="Vector: p773-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonafide, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      76 a      88 c      108 g      93 t
ORIGIN

```

```

Query Match      10.6%; Score 39; DB 10; Length 365;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 64 TGTGACGCTCGATTCACCTTCATATACACGCATG 102
      |||||||
DB 126 TGTGACGCTCGATTCACCTTCATATACACGCATG 164

```

```

RESULT 5      565 bp      mRNA      linear      EST 01-MAY-2001
LOCUS      BG686641
DEFINITION  BG686641 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4765555 5',
mRNA sequence.
ACCESSION  BG686641
VERSION    BG686641.1 GI:13918038
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens

```

```

REFERENCE 1 (bases 1 to 565)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LRCM1624 row: J column: 20
          High quality sequence start: 65.
          High quality sequence stop: 565.

```

```

FEATURES
source
location/Qualifiers
1..565
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4765555"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: B-cells; Vector: pOT7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

```

```

BASE COUNT      118 a      155 c      169 g      123 t
ORIGIN

```

```

Query Match      9.8%; Score 36; DB 12; Length 565;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      196 GTGAAAGACAGGTTACCATCTCCAGAGATGATTCA 231
LOCUS   BH319627
Db      306 GTGAAAGACAGGTTACCATCTCCAGAGATGATTCA 341

RESULT 6
AUTHORS BH319627
LOCUS   BH319627
DEFINITION CH230-119P2.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
ACCESSION BH319627
VERSION   BH319627.1 GI:17250341
KEYWORDS GSS.
SOURCE    Norway rat.
ORGANISM Rattus norvegicus
          Chordata: Craniata: Vertebrata: Euteleostomi:
          Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:
          Rattus.
          1 (bases 1 to 552)
REFERENCE Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
AUTHORS   ,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
          Jong,P. and Fraser,C.M.
          Rat BAC end sequences from library CHORI-230 EcoRI segment
          Unpublished (1999)
          Other_GSSs: CH230-119P2.TVB
          Contact: Shaying Zhao
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0200
          Fax: 301 838 0208
          Email: szhao@ligr.org
          Clones are derived from the rat BAC library CHORI-230
          (http://www.chori.org/bacpac/rat230.htm). For BAC library
          availability, please contact Pieter de Jong (pdejong@email.cho.org).
          Clones may be purchased from BACPAC Resources
          (http://www.chori.org/bacpac/or ering_information.htm). BAC end
          page: http://www.ligr.org/tbdb/bac_ends/rat/bac_end_intro.html
          Plate: 119 row: P column: 2
          Seq primer: SP6
          Class: BAC ends.
FEATURES
  source          Location/Qualifiers
    1..552
    /organism="Rattus norvegicus"
    /strain="BN/SSNhsd/MCW"
    /db_xref="taxon:10116"
    /clone_lib="CH230-119P2"
    /clone_1lb="CHORI-230 Segment 1"
    /sex="Female"
    /cell_type="Brain"
    /note="Vector: PTARBAC2.1; Site.1: EcoRI; Site.2: EcoRI;
    CHORI-230 Rat (BN/SSNhsd/MCW) BAC library produced by
    Pieter de Jong"

BASE COUNT      174 a      101 c      116 g      161 t

Query Match          9.5%; Score 35; DB 17; Length 552;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      19 TCTGTGTGAGAGATTGTGTCAGCCTAAAGGCTCATTT 53
LOCUS   BH319627
Db      26 TCTGTGTGAGAGATTGTGTCAGCCTAAAGGCTCATTT 60

RESULT 7
AUTHORS BG963642
LOCUS   BG963642
DEFINITION 602828443F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4983155 5',
ACCESSION BG963642
VERSION   BG963642.1 GI:14351279

```

```

KEYWORDS EST.
SOURCE    house mouse.
ORGANISM Mus musculus
          Chordata: Craniata: Vertebrata: Euteleostomi:
          Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
          1 (bases 1 to 666)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnl.gov
          Plate: L1AM10987 row: e column: 12
          High quality sequence stop: 658.
FEATURES
  source          Location/Qualifiers
    1..666
    /organism="Mus musculus"
    /strain="FVB/N"
    /db_xref="taxon:10090"
    /clone_lib="NCI_CGAP_Co24"
    /clone_1lb="NCI_CGAP_Co24"
    /lab_host="DH10B (T1 phage-resistant)"
    /note="Organ: colon; Vector: pCMV-SPORT6; Site.1: NotI;
    Site.2: SalI; Cloned unidirectionally. primer: Oligo dT.
    Average insert size 1.6 kb. Constructed by Life
    Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      157 a      186 c      159 g      164 t

Query Match          9.5%; Score 35; DB 13; Length 666;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      320 ACTATGCTATGACTACTGGGTCAGACGACCTCA 354
LOCUS   BG967007
Db      383 ACTATGCTATGACTACTGGGTCAGACGACCTCA 417

RESULT 8
AUTHORS BG967007
LOCUS   BG967007
DEFINITION 602834239F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4988744 5',
ACCESSION BG967007
VERSION   BG967007.1 GI:14354644
KEYWORDS EST.
SOURCE    house mouse.
ORGANISM Mus musculus
          Chordata: Craniata: Vertebrata: Euteleostomi:
          Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
          1 (bases 1 to 781)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnl.gov
          Plate: L1AM11001 row: n column: 09
          High quality sequence stop: 770.
FEATURES
  source          Location/Qualifiers
    1..781

```

/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4988744"  
/clone\_11b="NCI\_CGAP\_Co24"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPOrt6; Site:1: NotI;  
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI-CGAP Library."  
BASE COUNT 200 a 205 c 194 g 182 t  
ORIGIN

Query Match 9.5%; Score 35; DB 13; Length 781;  
Best Local Similarity 100.0%; Pred. No. 4.4e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 320 ACTATGCTATGACTACTGGGTCAGAACCTCA 354  
|||||  
Db 404 ACTATGCTATGACTACTGGGTCAGAACCTCA 438  
|||||

RESULT 9  
LOCUS B1455668 861 bp mRNA linear EST 21-AUG-2001  
DEFINITION 603173862F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:5253279 5',  
mRNA sequence.  
ACCESSION B1455668  
VERSION B1455668.1 GI:15246324  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 861)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM11639 row: 1 column: 16  
High quality sequence stop: 780.  
Location/Qualifiers  
1. 861  
/organism="Mus musculus"  
/strain="C57Bl/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5253279"  
/clone\_11b="NCI\_CGAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPOrt6; Site:1: SalI;  
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Rodin Humphreys,  
NIH"

BASE COUNT 205 a 239 c 211 g 206 t  
ORIGIN

Query Match 9.5%; Score 35; DB 13; Length 861;  
Best Local Similarity 100.0%; Pred. No. 4.5e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 320 ACTATGCTATGACTACTGGGTCAGAACCTCA 354  
|||||

Db 432 ACTATGCTATGACTACTGGGTCAGAACCTCA 466  
|||||

RESULT 10  
LOCUS B1102617 862 bp mRNA linear EST 26-JUN-2001  
DEFINITION 602888151F1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:5043444  
5', mRNA sequence.  
ACCESSION B1102617  
VERSION B1102617.1 GI:14553510  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 862)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1119 row: e column: 13  
High quality sequence stop: 849.  
Location/Qualifiers  
1. 862  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5043444"  
/clone\_11b="NCI\_CGAP\_Kid14"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: kidney; Vector: pCMV-SPOrt6; Site:1: NotI;  
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI-CGAP Library. 1"

BASE COUNT 211 a 231 c 209 g 211 t  
ORIGIN

Query Match 9.5%; Score 35; DB 13; Length 862;  
Best Local Similarity 100.0%; Pred. No. 4.5e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 320 ACTATGCTATGACTACTGGGTCAGAACCTCA 354  
|||||

Db 380 ACTATGCTATGACTACTGGGTCAGAACCTCA 414  
|||||

RESULT 11  
LOCUS BE309336 864 bp mRNA linear EST 26-OCT-2000  
DEFINITION 601093720F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3488309 5',  
mRNA sequence.  
ACCESSION BE309336  
VERSION BE309336.1 GI:9167366  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 864)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM8527 row: h column: 06  
High quality sequence stop: 645.  
Location/Qualifiers

FEATURES  
SOURCE

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_image="IMAGE:3488309"  
/clone\_id="NCI\_CGAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

BASE COUNT 171 a 260 c 222 g 211 t  
ORIGIN

Query Match 9.5%; Score 35; DB 10; Length 864;  
Best Local Similarity 100.0%; Pred. No. 4.5e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 320 ACTATGCTATGACTGCGGTCAAGACCTCA 354  
|||||  
Db 366 ACTATGCTATGACTGCGGTCAAGACCTCA 400

## RESULT 12

LOCUS B0921928 1010 bp mRNA linear EST 20-AUG-2002  
DEFINITION AGENCOURT-8952829 NCI\_CGAP\_Co24 Mus musculus CDNA clone  
IMAGE:6475284 5', mRNA sequence.  
B0921928  
B0921928.1 GI:22336959

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1010)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM4013 row: m column: 13  
High quality sequence stop: 608.  
Location/Qualifiers

FEATURES  
SOURCE

1. 1010  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone\_image="IMAGE:6475284"  
/clone\_id="NCI\_CGAP\_Co24"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 241 a 283 c 252 g 231 t 3 others  
ORIGIN

Query Match 9.5%; Score 35; DB 14; Length 1010;  
Best Local Similarity 100.0%; Pred. No. 4.8e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 320 ACTATGCTATGACTGCGGTCAAGACCTCA 354  
|||||  
Db 382 ACTATGCTATGACTGCGGTCAAGACCTCA 416

RESULT 13  
LOCUS B0266826 567 bp mRNA linear EST 07-MAY-2002  
DEFINITION NISC\_ff15c05.y1 NCI\_CGAP\_Mam5 Mus musculus CDNA clone IMAGE:265800  
5', mRNA sequence.  
B0266826  
B0266826.1 GI:20491891

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 567)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen/Robin Humphreys  
CDNA Library Preparation: Life Technologies  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
Info@image.llnl.gov  
MGI:1046572  
Plate: LLM7070 row: F column: 9  
Seq primer: M13RPI reverse primer (ABI).  
Location/Qualifiers

FEATURES  
SOURCE

1. 567  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_image="IMAGE:265800"  
/clone\_id="NCI\_CGAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

BASE COUNT 144 a 145 c 137 g 141 t  
ORIGIN

Query Match 9.2%; Score 34; DB 14; Length 567;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 321 CTATGCTATGACTGCGGTCAAGACCTCA 354  
|||||  
Db 398 CTATGCTATGACTGCGGTCAAGACCTCA 431

RESULT 14  
LOCUS BG966355 837 bp mRNA linear EST 12-JUN-2001



DEFINITION 602832843F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4987379 5',  
 mRNA sequence.  
 ACCESSION BG966355  
 VERSION BG966355.1 GI:14353992  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 837)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM10998 row: e column: 12  
 High quality sequence stop: 692.  
 Location/Qualifiers  
 1..837  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:4987379"  
 /clone\_11b="NCI\_CGAP\_Co24"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP library."  
 201 a 222 c 218 g 196 t

BASE COUNT  
 ORIGIN

Query Match 9.2%; Score 34; DB 13; Length 837;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;  
 Matches 34: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 CTATGCTATGACTGCGGTCAAGAACTCA 354  
 ||||||||||||||||||||||||||||||||  
 DB 446 CTATGCTATGACTGCGGTCAAGAACTCA 479

RESULT 15  
 BF578938 963 bp mRNA linear EST 12-DEC-2000  
 LOCUS 602095519F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4215533 5',  
 mRNA sequence.  
 ACCESSION BF578938  
 VERSION BF578938.1 GI:11652650  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 963)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM9791 row: e column: 06  
 High quality sequence start: 4  
 High quality sequence stop: 624.  
 Location/Qualifiers  
 1..963  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:4215533"  
 /clone\_11b="NCI\_CGAP\_Co24"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP library."  
 250 a 281 c 251 g 181 t

BASE COUNT  
 ORIGIN

Query Match 9.2%; Score 34; DB 12; Length 963;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;  
 Matches 34: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 CTATGCTATGACTGCGGTCAAGAACTCA 354  
 ||||||||||||||||||||||||||||||||  
 DB 122 CTATGCTATGACTGCGGTCAAGAACTCA 155

Search completed: November 27, 2002, 07:13:23  
 Job time : 1551.06 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:55:59 ; Search time 41.3581 Seconds  
(without alignments)  
2736.194 Million cell updates/sec

Title: US-09-893-615-86

Perfect score: 369  
Sequence: 1 GAAGTGATCTGCTGTGAGTC.....CCTCACTCAGCTCTCTCA 369

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents\_NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	132	35.8	330	3	US-08-767-128-23	Sequence 23, Appl
2	130	35.2	345	3	US-08-767-128-35	Sequence 35, Appl
3	80	21.7	357	1	US-08-442-542-1	Sequence 1, Appl
4	80	21.7	357	3	US-08-765-469-1	Sequence 1, Appl
5	80	21.7	1787	1	US-08-442-542-17	Sequence 17, Appl
6	80	21.7	1787	3	US-08-765-469-17	Sequence 17, Appl
7	36	9.8	351	2	US-08-752-844-36	Sequence 36, Appl
8	36	9.8	351	2	US-08-591-196-36	Sequence 36, Appl
9	36	9.8	738	1	US-08-197-834-6	Sequence 6, Appl
10	35	9.5	42	1	US-07-834-539A-41	Sequence 41, Appl
11	35	9.5	42	1	US-08-053-131-49	Sequence 49, Appl
12	35	9.5	42	1	US-08-645-641-49	Sequence 49, Appl
13	35	9.5	42	1	US-07-853-408B-49	Sequence 49, Appl
14	35	9.5	42	1	US-08-096-762-49	Sequence 49, Appl
15	35	9.5	42	2	US-08-800-353-41	Sequence 41, Appl
16	35	9.5	42	2	US-08-308-865-49	Sequence 49, Appl
17	35	9.5	42	4	US-09-042-353-212	Sequence 212, App
18	35	9.5	42	4	US-08-758-417A-60	Sequence 60, Appl
19	35	9.5	42	5	PCT-US92-06185-41	Sequence 41, Appl
20	35	9.5	42	5	PCT-US92-10983-49	Sequence 49, Appl
21	35	9.5	48	2	US-08-752-844-38	Sequence 38, Appl
22	35	9.5	48	2	US-08-591-196-38	Sequence 38, Appl
23	35	9.5	49	2	US-08-752-844-33	Sequence 33, Appl
24	35	9.5	49	2	US-08-591-196-33	Sequence 33, Appl
25	35	9.5	51	2	US-08-659-567-28	Sequence 28, Appl
26	35	9.5	57	2	US-08-752-844-30	Sequence 30, Appl
27	35	9.5	57	2	US-08-752-844-42	Sequence 42, Appl

28	35	9.5	57	2	US-08-591-196-30	Sequence 30, Appl
29	35	9.5	57	2	US-08-591-196-42	Sequence 42, Appl
30	35	9.5	346	3	US-08-881-037-13	Sequence 13, Appl
31	35	9.5	357	3	US-08-767-128-5	Sequence 5, Appl
32	35	9.5	360	1	US-08-447-422-14	Sequence 14, Appl
33	35	9.5	360	2	US-08-672-175A-1	Sequence 1, Appl
34	35	9.5	366	3	US-08-881-037-49	Sequence 49, Appl
35	35	9.5	372	3	US-08-767-128-37	Sequence 37, Appl
36	35	9.5	416	1	US-08-253-877C-7	Sequence 7, Appl
37	35	9.5	416	3	US-08-603-024-1	Sequence 1, Appl
38	35	9.5	417	2	US-08-452-164A-7	Sequence 7, Appl
39	35	9.5	420	4	US-08-943-136-3	Sequence 3, Appl
40	35	9.5	420	4	US-08-973-518-3	Sequence 3, Appl
41	35	9.5	711	6	5455030-8	Patent No. 5455030
42	35	9.5	732	1	US-08-230-843-1	Sequence 1, Appl
43	35	9.5	732	2	US-08-636-936-1	Sequence 1, Appl
44	35	9.5	741	6	5455030-10	Patent No. 5455030
45	35	9.5	754	1	US-07-958-140-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-767-128-23  
; Sequence 23, Application US/08767128  
; Patent No. 6111079  
; GENERAL INFORMATION:  
; APPLICANT: WYLLIE, DWANE E.  
; APPLICANT: LOPEZ, OSVALDO  
; APPLICANT: MURRAY, PETER JOSEPH  
; APPLICANT: GOEBEL, PETER JOSEPH  
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
; TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 6111079west Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/767,128  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09258  
; FILING DATE: 05-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/541,373  
; FILING DATE: 10-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/462,798  
; FILING DATE: 05-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carter, Charles G.  
; REGISTRATION NUMBER: 35,093  
; REFERENCE/DOCKET NUMBER: 8648.49USF1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/371-5278  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:  
LENGTH: 330 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1..330  
OTHER INFORMATION:  
US-08-767-128-23

Query Match 35.8%; Score 132; DB 3; Length 330;  
Best Local Similarity 99.5%; Pred. No. 4,4e-61;  
Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 CCAGGAAGGTTTGGATGGTTCCTCCATAGAAAGTAATAATTATGCAACA 180  
DB 106 CCAGGAAGGTTTGGATGGTTCCTCCATAGAAAGTAATAATTATGCAACA 165  
QY 181 TTTTATGCCGATTACAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAGCATG 240  
DB 166 TATTATGCCGATTACAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAGCATG 225  
QY 241 CTCTATCTGCAATGAACTTGAAGAACTGAGACACAGCCATGATTAAGTGTGAGA 300  
DB 226 CTCTATCTGCAATGAACTTGAAGAACTGAGACACAGCCATGATTAAGTGTGAGA 285  
QY 301 CGG 303  
DB 286 CGG 288

RESULT 2  
US-08-767-128-35  
Sequence 35, Application US/08767128  
Patent No. 6111079  
GENERAL INFORMATION:  
APPLICANT: WYLIE, DMANE E.  
APPLICANT: LOPEZ, OSVALDO  
APPLICANT: MURRAY, PETER JOSEPH  
APPLICANT: GOEBEL, PETER  
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 6111079west Center, 90 South Seventh st  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,128  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09258  
FILING DATE: 05-JUN-1996  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/541,373  
FILING DATE: 10-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/462,798  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Carter, Charles G.  
REGISTRATION NUMBER: 35,093  
REFERENCE/DOCKET NUMBER: 8648.49USF1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612/371-5278  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 345 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1..345  
OTHER INFORMATION:  
US-08-767-128-35

Query Match 35.2%; Score 130; DB 3; Length 345;  
Best Local Similarity 99.4%; Pred. No. 5,2e-60;  
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 123 AGGAAGGTTTGGATGGTTCCTCCATAGAAAGTAATAATTATGCAACATT 182  
DB 123 AGGAAGGTTTGGATGGTTCCTCCATAGAAAGTAATAATTATGCAACATA 182  
QY 183 TTATGCCGATTACAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAGCATGCT 242  
DB 183 TTATGCCGATTACAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAGCATGCT 242  
QY 243 CTATCTGCAATGAACTTGAAGAACTGAGACACAGCCATGATTAAGTGTGAGA 302  
DB 243 CTATCTGCAATGAACTTGAAGAACTGAGACACAGCCATGATTAAGTGTGAGA 302  
QY 303 G 303  
DB 303 G 303

RESULT 3  
US-08-442-542-1  
Sequence 1, Application US/08442542  
Patent No. 5686600  
GENERAL INFORMATION:  
APPLICANT: Kozel, Nadine B.  
APPLICANT: Kozel, Michael G.  
TITLE OF INVENTION: Antibodies which bind to Insect Gut  
TITLE OF INVENTION: Proteins and their Use  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,542  
FILING DATE: 16-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/267,641  
FILING DATE: 28-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 357 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..357  
OTHER INFORMATION: /note="3B1 heavy chain variable"  
OTHER INFORMATION: region from pcib4613"  
US-08-442-542-1

Query Match 21.7%; Score 80; DB 1; Length 357;  
Best Local Similarity 100.0%; Pred. No. 2.4e-33;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GGAGCTGTGGAGATTGGTGCACCTTAAAGGTCATTGAAACTCTCATGTGCAGCCTC 74  
DB 15 GGAGCTGTGGAGATTGGTGCACCTTAAAGGTCATTGAAACTCTCATGTGCAGCCTC 74  
QY 75 TGGATTCACTTCATAACT 94  
DB 75 TGGATTCACTTCATAACT 94

RESULT 4  
US-08-765-469-1  
Sequence 1, Application US/08765469  
Patent No. 6069301  
GENERAL INFORMATION:  
APPLICANT: Carozzi, Nadine B.  
APPLICANT: Kozziel, Michael G.  
TITLE OF INVENTION: Antibodies which bind to Insect Gut  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/765,469  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/267,641  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 357 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..357  
OTHER INFORMATION: /note="3B1 heavy chain variable"  
OTHER INFORMATION: region from pcib4613"  
US-08-765-469-1

Query Match 21.7%; Score 80; DB 3; Length 357;  
Best Local Similarity 100.0%; Pred. No. 2.4e-33;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GGAGCTGTGGAGATTGGTGCACCTTAAAGGTCATTGAAACTCTCATGTGCAGCCTC 74  
DB 15 GGAGCTGTGGAGATTGGTGCACCTTAAAGGTCATTGAAACTCTCATGTGCAGCCTC 74  
QY 75 TGGATTCACTTCATAACT 94  
DB 75 TGGATTCACTTCATAACT 94

RESULT 5  
US-08-442-542-17  
Sequence 17, Application US/08442542  
Patent No. 5686600  
GENERAL INFORMATION:  
APPLICANT: Carozzi, Nadine B.  
APPLICANT: Kozziel, Michael G.  
TITLE OF INVENTION: Antibodies which bind to Insect Gut  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,542  
FILING DATE: 16-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/267,641  
FILING DATE: 28-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1797 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1797  
OTHER INFORMATION: /note="3B1 single chain antibody"  
OTHER INFORMATION: from PCIB4631"  
US-08-442-542-17

Query Match  
Best Local Similarity 100.0%; Pred. No. 2.6e-33;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GGAGTCGTGGAGAGATTGGTGCAGCCTAAAGGCTCATTTGAACCTCATGTGCAGCCTC 74  
|||||  
DB 477 GGAGTCGTGGAGAGATTGGTGCAGCCTAAAGGCTCATTTGAACCTCATGTGCAGCCTC 536  
QY 75 TGGATTCACCTTCATTAATCT 94  
|||||  
DB 537 TGGATTCACCTTCATTAATCT 556

## RESULT 6

US-08-765-469-117  
Sequence 17, Application US/08765469  
Patent No. 6069301  
GENERAL INFORMATION:  
APPLICANT: Carozzi, Nadine B.  
APPLICANT: Koziel, Michael G.  
TITLE OF INVENTION: Antibodies which bind to insect gut  
TITLE OF INVENTION: Proteins and their use  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/765,469  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/267,641  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1797 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1797  
OTHER INFORMATION: /note="3B1 single chain antibody"  
OTHER INFORMATION: from PCIB4631"  
US-08-765-469-117

Query Match  
Best Local Similarity 100.0%; Pred. No. 2.6e-33;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GGAGTCGTGGAGAGATTGGTGCAGCCTAAAGGCTCATTTGAACCTCATGTGCAGCCTC 74  
|||||  
DB 477 GGAGTCGTGGAGAGATTGGTGCAGCCTAAAGGCTCATTTGAACCTCATGTGCAGCCTC 536  
QY 75 TGGATTCACCTTCATTAATCT 94  
|||||  
DB 537 TGGATTCACCTTCATTAATCT 556

## RESULT 7

US-08-752-844-36  
Sequence 36, Application US/08752844  
Patent No. 5935821  
GENERAL INFORMATION:  
APPLICANT: Chatterjee, Malaya  
APPLICANT: Foon, Kenneth A.  
TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE  
TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,844  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Schiff, J. Michael  
REGISTRATION NUMBER: 40,253  
REFERENCE/DOCKET NUMBER: 30414-20002.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-752-844-36

Query Match  
Best Local Similarity 100.0%; Pred. No. 6.9e-10;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 GACTATGCTATGACTACTAGGGGTCAAGGAACTCA 354  
|||||  
DB 304 GACTATGCTATGACTACTAGGGGTCAAGGAACTCA 339

## RESULT 8

US-08-591-196-36  
Sequence 36, Application US/08591196  
Patent No. 5977316  
GENERAL INFORMATION:  
APPLICANT: Chatterjee, Malaya  
APPLICANT: Foon, Kenneth A.  
TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE

TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/591,196  
FILING DATE: 16-JAN-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Schifff, J. Michael  
REGISTRATION NUMBER: 40,253  
REFERENCE/DOCKET NUMBER: 30414-20002.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ. ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-591-196-36

Query Match 9.8%; Score 36; DB 2; Length 351;  
Best Local Similarity 100.0%; Pred. No. 6.9e-10;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 GACTATGCTATGGACTACTGGGTCGAAGAACTCA 354  
|||||  
DB 304 GACTATGCTATGGACTACTGGGTCGAAGAACTCA 339

RESULT 9  
US-08-197-834-6  
Sequence 6, Application US/08197834  
Patent No. 5639455  
GENERAL INFORMATION:  
APPLICANT: SHIMAMURA, TOSHIRO  
APPLICANT: NAKAZAWA, HARUMI  
APPLICANT: HAMURO, JUNJI  
TITLE OF INVENTION: IMMUNOSUPPRESSANT  
NUMBER OF SEQUENCES: 7  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/197,834  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 028173/1993

FILING DATE: 17-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NO. 5639455man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 10-661-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 738 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Other nucleic acid;  
DESCRIPTION: DNA (synthetic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..738  
US-08-197-834-6

Query Match 9.8%; Score 36; DB 1; Length 738;  
Best Local Similarity 100.0%; Pred. No. 7.2e-10;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 GACTATGCTATGGACTACTGGGTCGAAGAACTCA 354  
|||||  
DB 688 GACTATGCTATGGACTACTGGGTCGAAGAACTCA 723

RESULT 10  
US-07-834-539A-41  
Sequence 41, Application US/07834539A

GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
APPLICANT: Kay, Robert M.  
TITLE OF INVENTION: Transgenic Non-Human Animals Capable of  
Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 77  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/834,539A  
FILING DATE: 1992-02-05  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ. ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-834-539A-41

Query Match 9.5%; Score 35; DB 1; Length 42;  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/834,539A  
FILING DATE: 1992-02-05  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ. ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-834-539A-41

Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 320 ACTATGCTATGACTACTGGGTCAAGAACCTCA 354  
Db 1 ACTATGCTATGACTACTGGGTCAAGAACCTCA 35

## RESULT 11

US-08-053-131-49  
Sequence 49, Application US/08053131  
Patent No. 5661016  
GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
APPLICANT: Kay, Robert M.  
TITLE OF INVENTION: Transgenic No. 5661016-Human Animals for  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Hourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 200  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/053,131  
FILING DATE: 26-APR-1993  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: US 07/990,860  
FILING DATE: 16-DEC-1992  
PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: US 07/810,279  
FILING DATE: 17-DEC-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/853,408  
FILING DATE: 18-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-9-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (primer)  
US-08-053-131-49

Query Match 9.5%; Score 35; DB 1; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 320 ACTATGCTATGACTACTGGGTCAAGAACCTCA 354  
Db 1 ACTATGCTATGACTACTGGGTCAAGAACCTCA 35

## RESULT 12

US-08-645-641-49  
Sequence 49, Application US/08645641  
Patent No. 5719032  
GENERAL INFORMATION:

APPLICANT: Lonberg, Nils  
APPLICANT: Kay, Robert M.  
TITLE OF INVENTION: Transgenic No. 5719032-Human Animals for  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 150  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/645,641  
FILING DATE: 20-MAY-1996  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: US 07/904,068  
FILING DATE: 23-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-000913  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (primer)  
US-08-645-641-49

Query Match 9.5%; Score 35; DB 1; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 320 ACTATGCTATGACTACTGGGTCAAGAACCTCA 354  
Db 1 ACTATGCTATGACTACTGGGTCAAGAACCTCA 35

## RESULT 13

US-07-853-408B-49  
Sequence 49, Application US/07853408B  
Patent No. 5789650  
GENERAL INFORMATION:

APPLICANT: Lonberg, Nils  
APPLICANT: Kay, Robert M.  
TITLE OF INVENTION: Transgenic No. 5789650-Human Animals for  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 150  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/853,408B



FILING DATE: 19920318  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ. ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (primer)  
US-07-853-4088-49

Query Match  
Best Local Similarity 100.0%; Score: 35; DB 1; Length 42;  
Pred. No. 2.1e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 ACTATGCTATGACTACTGGGCTCAGAGAACTCTCA 354  
DB 1 ACTATGCTATGACTACTGGGCTCAGAGAACTCTCA 35

RESULT 14  
US-08-096-762-49  
Sequence 49, Application US/08096762  
Patent No. 5814318  
GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
TITLE OF INVENTION: Transgenic No. 5814318-Human Animals for  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 210  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Knourle and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 200  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/096,762  
FILING DATE: 22-JUL-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/053,131  
FILING DATE: 26-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/990,860  
FILING DATE: 16-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/904,068  
FILING DATE: 23-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/853,408  
FILING DATE: 18-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/810,279  
FILING DATE: 17-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-9-4  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ. ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (primer)  
US-08-096-762-49

Query Match  
Best Local Similarity 100.0%; Score: 35; DB 1; Length 42;  
Pred. No. 2.1e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 ACTATGCTATGACTACTGGGCTCAGAGAACTCTCA 354  
DB 1 ACTATGCTATGACTACTGGGCTCAGAGAACTCTCA 35

RESULT 15  
US-08-800-353-41  
Sequence 41, Application US/08800353  
Patent No. 5874299  
GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
TITLE OF INVENTION: Transgenic No. 5874299-Human Animals Capable of  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/800,353  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/834,539  
FILING DATE: 1992-02-05  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ. ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-800-353-41

Query Match  
Best Local Similarity 100.0%; Score: 35; DB 2; Length 42;  
Pred. No. 2.1e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 ACTATGCTATGACTACTGGGCTCAGAGAACTCTCA 354  
DB 1 ACTATGCTATGACTACTGGGCTCAGAGAACTCTCA 35

Search completed: November 27, 2002, 05:39:23  
Job time : 49.3581 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 05:34:00 : Search time 41.3581 Seconds

(without alignments)  
3436.030 Million cell updates/sec

Title: US-09-893-615-86

Perfect score: 369  
Sequence: 1 GAAGGATGCTGTGTGAGTC.....CCTCACTACGCTCTCTCA 369

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 341543 seqs, 192557720 residues

Word size : 0

Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	100.0	369	10	US-09-893-615-86
2	101	27.4	466	10	US-09-881-823-15
3	58	15.7	443	10	US-09-840-459-96
4	58	15.7	443	10	US-09-840-459-96
5	40	10.8	357	10	US-09-840-459-97
6	40	10.8	357	10	US-09-840-459-97
7	36	9.8	478	10	US-09-840-459-911
8	35	9.5	51	7	US-10-040-739-911
9	35	9.5	417	7	US-08-779-784-13
10	35	9.5	420	10	US-09-007-093-3
11	35	9.5	466	10	US-09-881-823-11
12	34	9.2	354	10	US-09-949-559-105
13	34	9.2	414	9	US-10-146-305-6
14	33	8.9	736	12	US-10-006-773-12
15	32	8.7	418	10	US-09-753-436-42
16	30	8.1	360	10	US-09-910-059-10
17	30	8.1	765	10	US-09-910-059-18
18	29	7.9	42	9	US-09-252-150-44
19	29	7.9	354	10	US-09-905-243-71

20	29	7.9	396	10	US-09-811-737-12	Sequence 12, Appl
21	29	7.9	396	10	US-09-811-737-13	Sequence 13, Appl
22	29	7.9	420	10	US-09-286-240-3	Sequence 3, Appl
23	29	7.9	782	10	US-09-811-737-21	Sequence 21, Appl
24	29	7.9	782	10	US-09-811-737-22	Sequence 22, Appl
25	28	7.6	372	9	US-09-879-813-87	Sequence 87, Appl
26	28	7.6	454	10	US-09-881-823-3	Sequence 3, Appl
27	28	7.6	518	10	US-09-881-823-17	Sequence 17, Appl
28	27	7.3	351	8	US-08-790-540A-5	Sequence 5, Appl
29	27	7.3	351	8	US-08-791-391A-5	Sequence 5, Appl
30	26	7.0	422	10	US-09-753-436-77	Sequence 77, Appl
31	26	7.0	458	10	US-09-990-205-3	Sequence 3, Appl
32	26	7.0	461	10	US-09-861-294-3	Sequence 3, Appl
33	26	7.0	690	10	US-09-995-693-4	Sequence 4, Appl
34	26	7.0	714	8	US-08-940-544-4	Sequence 4, Appl
35	26	7.0	1598	10	US-09-822-849A-103	Sequence 103, App
36	26	7.0	1617	10	US-09-822-830A-571	Sequence 571, App
37	25	6.8	132	9	US-09-423-800-25	Sequence 25, App
38	25	6.8	308	10	US-09-864-761-27741	Sequence 27741, A
39	25	6.8	311	10	US-09-864-761-31233	Sequence 31233, A
40	25	6.8	335	10	US-09-828-708-111	Sequence 111, App
41	25	6.8	331	8	US-08-790-540A-1	Sequence 1, Appl
42	25	6.8	331	8	US-08-791-391A-1	Sequence 1, Appl
43	25	6.8	351	9	US-10-091-236-7	Sequence 7, Appl
44	25	6.8	354	10	US-09-798-058-1	Sequence 1, Appl
45	25	6.8	397	9	US-10-046-935-2106	Sequence 2106, Ap

#### ALIGNMENTS

RESULT 1  
US-09-893-615-86  
Sequence 86, Application US/09893615  
Patent No. US20020082395A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Gerald W.  
Schuman, Richard F.  
Wong, Hing  
Stinson, Jeffrey L.  
TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GI  
POSITIVE BACTERIA  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARROW, GARRETT &  
DUNN, LLP  
STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/893,615  
FILING DATE: 29-Jun-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: Einaudi, Carol P  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 04995.0041-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..369  
SEQUENCE DESCRIPTION: SEQ ID NO: 86  
US-09-893-615-86

Query Match 100.0%; Score 369; DB 10; Length 369;  
Best Local Similarity 100.0%; Pred. No. 1,1e-188;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAAGTATGCTGCTGAGCTGCTGCTGAGAGATTGTCAGCCCTAAAGGCTCATTTGAACCTC 60  
Db 1 GAAGTATGCTGCTGAGCTGCTGCTGAGAGATTGTCAGCCCTAAAGGCTCATTTGAACCTC 60  
Oy 61 TCATGTGAGCCCTGAGATTACCTTCATAAATAGCCCAATGAGTGGCTCCGCAAGCT 120  
Db 61 TCATGTGAGCCCTGAGATTACCTTCATAAATAGCCCAATGAGTGGCTCCGCAAGCT 120  
Oy 121 CCAGGAAGAGGTTTGAATGGGTTCTGCAATAGAAATAAATATTTATTCACACA 180  
Db 121 CCAGGAAGAGGTTTGAATGGGTTCTGCAATAGAAATAAATATTTATTCACACA 180  
Oy 181 TTTTATGCCGATTACGTAAAGACAGGTTTCAACCATCTCCAGAGATGATTCACAAAGCATG 240  
Db 181 TTTTATGCCGATTACGTAAAGACAGGTTTCAACCATCTCCAGAGATGATTCACAAAGCATG 240  
Oy 241 CTCTATGCAATGAAACACTTGAACACTGGAGACAGCCCATGTTATTCGTGTGAGA 300  
Db 241 CTCTATGCAATGAAACACTTGAACACTGGAGACAGCCCATGTTATTCGTGTGAGA 300  
Oy 301 CGGGGGGCTTACGAGTATGACTATGCTATGACTACTGAGGCTCAAGAACTCCTACTCACC 360  
Db 301 CGGGGGGCTTACGAGTATGACTATGACTACTGAGGCTCAAGAACTCCTACTCACC 360  
Oy 361 GTCTCCTCA 369  
Db 361 GTCTCCTCA 369

RESULT 2  
US-09-881-823-15  
Sequence 15, Application US/09881823  
Patent No. US20020068066A1  
GENERAL INFORMATION:  
APPLICANT: SHI, WENYUAN  
APPLICANT: ANDERSON, MAXWELL  
APPLICANT: MORRISON, SHERIE  
APPLICANT: TRINH, RYAN  
APPLICANT: WIMS, LETITIA  
APPLICANT: CHEN, LI  
TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries  
FILE REFERENCE: 22851-032  
CURRENT APPLICATION NUMBER: US/09/881,823  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 07/378,577  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 15  
LENGTH: 466  
TYPE: DNA  
ORGANISM: Murine  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (14)..(442)  
US-09-881-823-15

Query Match 27.4%; Score 101; DB 10; Length 466;  
Best Local Similarity 100.0%; Pred. No. 1e-44;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 200 AAGACAGTTACCATCTCCAGAGATGATTCACAAAGCATGCTATCTGCAATGAACA 259

Db 270 AAGACAGTTACCATCTCCAGAGATGATTCACAAAGCATGCTATCTGCAATGAACA 329  
Oy 260 ACTTGAACCTGAGACACACAGCATGATTTACTGTGTGAGA 300  
Db 330 ACTTGAACCTGAGACACACAGCATGATTTACTGTGTGAGA 370

RESULT 3  
US-09-840-459-96  
Sequence 96, Application US/09840459  
Patent No. US20020150576A1  
GENERAL INFORMATION:  
APPLICANT: Larosa, Gregory J.  
APPLICANT: Horvath, Christopher  
APPLICANT: Newman, Walter  
APPLICANT: Jones, S. Tarran  
APPLICANT: O'Brien, Siobhan H.  
TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
FILE REFERENCE: 1855.1052-012  
CURRENT APPLICATION NUMBER: US/09/840,459  
PRIOR FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: PCT/US01/03537  
PRIOR FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: 09/497,625  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 09/359,193  
PRIOR FILING DATE: 1998-07-23  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 96  
LENGTH: 443  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-840-459-96

Query Match 15.7%; Score 58; DB 10; Length 443;  
Best Local Similarity 100.0%; Pred. No. 1.3e-21;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 25 GGAGATTGTGACAGCCTAAAGGCTGATTTGAACCTGATGTCAGCCCTTGATTTCA 82  
Db 82 GGAGATTGTGACAGCCTAAAGGCTGATTTGAACCTGATGTCAGCCCTTGATTTCA 139  
RESULT 4  
US-09-840-459-99/c  
Sequence 99, Application US/09840459  
Patent No. US20020150576A1  
GENERAL INFORMATION:  
APPLICANT: Larosa, Gregory J.  
APPLICANT: Horvath, Christopher  
APPLICANT: Newman, Walter  
APPLICANT: Jones, S. Tarran  
APPLICANT: O'Brien, Siobhan H.  
APPLICANT: O'Keefe, Theresa  
TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
FILE REFERENCE: 1855.1052-012  
CURRENT APPLICATION NUMBER: US/09/840,459  
PRIOR FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: PCT/US01/03537  
PRIOR FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: 09/497,625  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 09/359,193  
PRIOR FILING DATE: 1998-07-23  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 99  
LENGTH: 443  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-840-459-99/c

NUMBER OF SEQ ID NOS: 107  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 99  
LENGTH: 443  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-840-459-99

Query Match  
Best Local Similarity 100.0%; Score 58; DB 10; Length 443;  
Pred. No. 1.3e-21;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GGAGGATGTGTCAGCCTTAAGGGCTGATGAACCTCTCATGTGCAGCTCTGGATTCA 82  
|||||  
Db 362 GGAGGATGTGTCAGCCTTAAGGGCTGATGAACCTCTCATGTGCAGCCTCTGGATTCA 305

## RESULT 5

US-09-840-459-97  
Sequence 97, Application US/09840459  
Patent No. US20020150576A1

## GENERAL INFORMATION:

APPLICANT: Larosa, Gregory J.  
APPLICANT: Horvath, Christopher  
APPLICANT: Newman, Walter  
APPLICANT: Jones, S. Tarran  
APPLICANT: O'Brien, Siobhan H.  
APPLICANT: O'Keefe, Theresa

TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
METHODS OF USE THEREFOR

FILE REFERENCE: 1855.1052-012

CURRENT APPLICATION NUMBER: US/09/840,459

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: PCT/US01/03537

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/497,625

PRIOR FILING DATE: 2000-02-03

PRIOR APPLICATION NUMBER: 09/359,193

PRIOR FILING DATE: 1999-07-22

PRIOR APPLICATION NUMBER: 09/121,781

PRIOR FILING DATE: 1998-07-23

NUMBER OF SEQ ID NOS: 107

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 97

LENGTH: 357

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Humanized heavy chain

Query Match  
Best Local Similarity 10.8%; Score 40; DB 10; Length 357;  
Pred. No. 5.8e-12;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 TGGGTCGCCAGGCTCCAGGAAGGTTGGATGGCTTG 145  
|||||  
Db 106 TGGGTCGCCAGGCTCCAGGAAGGTTGGATGGCTTG 145

## RESULT 6

US-09-840-459-103/C  
Sequence 103, Application US/09840459  
Patent No. US20020150576A1

## GENERAL INFORMATION:

APPLICANT: Larosa, Gregory J.  
APPLICANT: Horvath, Christopher  
APPLICANT: Newman, Walter  
APPLICANT: Jones, S. Tarran  
APPLICANT: O'Brien, Siobhan H.  
APPLICANT: O'Keefe, Theresa

TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND

METHODS OF USE THEREFOR

FILE REFERENCE: 1855.1052-012

CURRENT APPLICATION NUMBER: US/09/840,459

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: PCT/US01/03537

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/497,625

PRIOR FILING DATE: 2000-02-03

PRIOR APPLICATION NUMBER: 09/359,193

FILE REFERENCE: 1855.1052-012  
CURRENT APPLICATION NUMBER: US/09/840,459  
CURRENT FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: PCT/US01/03537  
PRIOR FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: 09/497,625  
PRIOR FILING DATE: 2000-02-03  
PRIOR APPLICATION NUMBER: 09/359,193  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 09/121,781  
PRIOR FILING DATE: 1998-07-23  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 103  
LENGTH: 357  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Humanized heavy chain  
US-09-840-459-103

Query Match  
Best Local Similarity 10.8%; Score 40; DB 10; Length 357;  
Pred. No. 5.8e-12;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 TGGGTCGCCAGGCTCCAGGAAGGTTGGATGGCTTG 145  
|||||  
Db 252 TGGGTCGCCAGGCTCCAGGAAGGTTGGATGGCTTG 213

## RESULT 7

US-10-040-739-911  
Sequence 911, Application US/10040739  
Patent No. US20020173635A1

## GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John

Lavallee, Edward

Racie, Lisa

Merberg, David

Treacy, Maurice

Spaulding, Vikki

TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS

NUMBER OF SEQUENCES: 1519

CORRESPONDENCE ADDRESS:

ADDRESSER: Genetics Institute, Inc.

STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/040,739

FILING DATE: 07-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/036,520

FILING DATE: 03-JUN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.

REGISTRATION NUMBER: 32,724

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8224

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 911:

SEQUENCE CHARACTERISTICS:

LENGTH: 478 base pairs

TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 911  
US-10-040-739-911

Query Match 9.8%; Score 36; DB 9; Length 478;  
Best Local Similarity 100.0%; Pred. No. 8.3e-10;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 GACTATGCTATGACTACTGGGGTCACAGAACCTCA 354  
DB 413 GACTATGCTATGACTACTGGGGTCACAGAACCTCA 448

RESULT 8  
US-08-779-784-13  
Sequence 13, Application US/08779784  
Patent No. US20020164325A1  
GENERAL INFORMATION:  
APPLICANT: Rodriguez, Moses  
APPLICANT: Miller, David J.  
APPLICANT: Asakura, Kunihiko  
TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM  
TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
STREET: Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/779,784  
FILING DATE: 07-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/692,084  
FILING DATE: 08-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/236,520  
FILING DATE: 29-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 51 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
US-08-779-784-13

Query Match 9.5%; Score 35; DB 7; Length 51;  
Best Local Similarity 100.0%; Pred. No. 2.5e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 ACTATGCTATGACTACTGGGGTCACAGAACCTCA 354  
DB 2 ACTATGCTATGACTACTGGGGTCACAGAACCTCA 36

RESULT 9  
US-08-779-784-7  
Sequence 7, Application US/08779784  
Patent No. US20020164325A1  
GENERAL INFORMATION:  
APPLICANT: Rodriguez, Moses  
APPLICANT: Miller, David J.  
APPLICANT: Asakura, Kunihiko  
TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM  
TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
STREET: Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/779,784  
FILING DATE: 07-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/692,084  
FILING DATE: 08-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/236,520  
FILING DATE: 29-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 417 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
US-08-779-784-7

Query Match 9.5%; Score 35; DB 7; Length 417;  
Best Local Similarity 100.0%; Pred. No. 2.8e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 ACTATGCTATGACTACTGGGGTCACAGAACCTCA 354  
DB 368 ACTATGCTATGACTACTGGGGTCACAGAACCTCA 402

RESULT 10  
US-09-007-093-3  
Sequence 3, Application US/09007093  
Patent No. US20020025315A1  
GENERAL INFORMATION:  
APPLICANT: Anand, Naveen N  
APPLICANT: Barber, Brian H  
APPLICANT: Cates, George A  
APPLICANT: Caterini, Judith E

APPLICANT: Klein, Michel H  
TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF  
NUMBER OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/007,093  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/483,576  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-007-093-3

Query Match  
Best Local Similarity 9.5%; Score 35; DB 10; Length 420;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 320 ACTATGCTATGACTACTGGGTCAGAGACCTCA 354  
|||||  
DB 371 ACTATGCTATGACTACTGGGTCAGAGACCTCA 405

RESULT 11  
US-09-881-823-11  
Sequence 11, Application US/09881823  
Patent No. US20020068066A1  
GENERAL INFORMATION:  
APPLICANT: SHI, WENTUAN  
APPLICANT: ANDERSON, MAXWELL  
APPLICANT: MORRISON, SHERIE  
APPLICANT: TRINH, RYAN  
APPLICANT: WIMS, LETITIA  
APPLICANT: CHEN, LI  
TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries  
FILE REFERENCE: 22851-032  
CURRENT APPLICATION NUMBER: US/09/881,823  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 07/378,577  
PRIOR FILING DATE: 1999-08-20  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 11  
LENGTH: 466  
TYPE: DNA  
ORGANISM: Murline  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (11)..(442)

US-09-881-823-11  
Query Match  
Best Local Similarity 9.5%; Score 35; DB 10; Length 466;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 320 ACTATGCTATGACTACTGGGTCAGAGACCTCA 354  
|||||  
DB 387 ACTATGCTATGACTACTGGGTCAGAGACCTCA 421

RESULT 12  
US-09-949-559-105  
Sequence 105, Application US/09949559  
Patent No. US20020151682A1  
GENERAL INFORMATION:  
APPLICANT: Athwal, Diljeet Singh  
APPLICANT: Brown, Derek Thomas  
APPLICANT: Weir, Andrew Neil Charles  
APPLICANT: Popplewell, Andrew George  
APPLICANT: Chapman, Andrew Paul  
APPLICANT: King, David John  
TITLE OF INVENTION: Biological Products  
FILE REFERENCE: Carp-0095  
CURRENT APPLICATION NUMBER: US/09/949,559  
CURRENT FILING DATE: 2001-12-20  
PRIOR APPLICATION NUMBER: 0013810.7GB  
PRIOR FILING DATE: 2000-06-06  
PRIOR APPLICATION NUMBER: 09/875,221  
PRIOR FILING DATE: 2001-06-06  
NUMBER OF SEQ ID NOS: 130  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 105  
LENGTH: 354  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: hTNF40 heavy chain variable domain  
NAME/KEY: CDS  
LOCATION: (1)..(354)  
OTHER INFORMATION:  
US-09-949-559-105

Query Match  
Best Local Similarity 9.2%; Score 34; DB 10; Length 354;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 321 CTATGCTATGACTACTGGGTCAGAGACCTCA 354  
|||||  
DB 306 CTATGCTATGACTACTGGGTCAGAGACCTCA 339

RESULT 13  
US-10-146-305-6  
Sequence 6, Application US/10146305  
Patent No. US20020173035A1  
GENERAL INFORMATION:  
APPLICANT: YIHAN CORPORATION  
TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV  
FILE REFERENCE: OV17440  
CURRENT APPLICATION NUMBER: US/10/146,305  
CURRENT FILING DATE: 2002-05-15  
PRIOR APPLICATION NUMBER: KR 10-2001-26634  
PRIOR FILING DATE: 2001-05-16  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 6  
LENGTH: 414  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-10-146-305-6

Query Match 9.2%; Score 34; DB 9; Length 414;  
Best Local Similarity 100.0%; Pred. No. 9.8e-09;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 321 CTATGCTATGACTCTGGGGTCAGGAACCTCA 354  
|||||

Db 318 CTATGCTATGACTCTGGGGTCAGGAACCTCA 351

RESULT 14  
US-10-006-773-12  
Sequence 12, Application US/10006773  
Patent No. US20020132983A1  
GENERAL INFORMATION:  
APPLICANT: Jungmans, Richard P.  
TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-  
FILE REFERENCE: 003  
CURRENT APPLICATION NUMBER: US/10/006,773  
CURRENT FILING DATE: 2001-12-10  
PRIOR APPLICATION NUMBER: 60/250,089  
PRIOR FILING DATE: 2000-11-30  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 736  
TYPE: DNA  
ORGANISM: Mus sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (14)..(430)  
OTHER INFORMATION: 4D4 Heavy chain V region, plus leader  
US-10-006-773-12

Query Match 8.9%; Score 33; DB 12; Length 736;  
Best Local Similarity 100.0%; Pred. No. 3.5e-08;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 TATGCTATGACTCTGGGGTCAGGAACCTCA 354  
|||||

Db 383 TATGCTATGACTCTGGGGTCAGGAACCTCA 415

RESULT 15  
US-09-753-436-42  
Sequence 42, Application US/09753436  
Patent No. US20010029293A1  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemary  
TITLE OF INVENTION: ICAM-Related Materials and Methods  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/753,436  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/382,289  
FILING DATE:  
APPLICATION NUMBER: US 08/487,113  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/286,754  
FILING DATE: 05-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Joseph A., Jr.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 33282  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 418 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-753-436-42

Query Match 8.7%; Score 32; DB 10; Length 418;  
Best Local Similarity 100.0%; Pred. No. 1.2e-07;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 323 ATGCTATGACTCTGGGGTCAGGAACCTCA 354  
|||||

Db 371 ATGCTATGACTCTGGGGTCAGGAACCTCA 402

Search completed: November 27, 2002, 07:14:51  
Job time : 50.3581 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:19:05 : Search time 43.6911 Seconds  
(without alignments)  
375.129 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 123  
Sequence: 1 EWMVSESGGLVOPKSLKLT.....SGIDYAMDYWGQGSTLVSS 123

Scoring table:

OLIGO  
Gapop 60.0 , Capext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	32.5	110	18	AAW01587
2	40	32.5	115	18	AAW01593
3	38	30.9	100	22	AAE06972
4	36	29.3	333	21	AAV70111
5	32	26.0	32	20	AAW94738
6	30	24.4	30	20	AAW94735
7	28	22.8	119	17	AAW90829
8	28	22.8	599	17	AAW90837
9	24	19.5	101	22	AAE06971
10	24	19.5	117	22	AAW09919

11	24	19.5	117	22	AAE06947	Murine ID9 antibody
12	24	19.5	117	23	AAO14978	Mouse heavy chain
13	24	19.5	148	22	AAE07033	Murine antibody ID
14	23	18.7	23	20	AAW94739	Anti-Staph (HAY) 9
15	21	17.1	117	22	AAU09928	Humanised ID9 heav
16	21	17.1	117	22	AAU09929	Humanised ID9 heav
17	21	17.1	117	22	AAU09930	Humanised ID9 heav
18	21	17.1	117	22	AAE06956	Humanised murine 1
19	21	17.1	117	22	AAE06957	Humanised murine 1
20	21	17.1	117	23	AAO14982	Humanised murine h
21	21	17.1	118	23	AAO14983	Humanised murine h
22	21	17.1	125	21	AAAB10022	H. pylori 26 kDa p
23	19	15.4	19	20	AAW94736	Anti-Staph (HAY) 9
24	19	15.4	19	20	AAW94737	Anti-Staph (HAY) 9
25	18	14.6	142	14	AAAR31534	MH4H7 Mab heavy ch
26	18	14.6	253	23	ABP45326	Human Blys binding
27	17	13.8	62	16	AAW76969	HSV-neutralising a
28	17	13.8	117	22	AAU09927	Humanised ID9 heav
29	17	13.8	117	22	AAE06954	Humanised murine 1
30	17	13.8	117	22	AAE06955	Humanised murine 1
31	17	13.8	117	23	AAO14980	Humanised murine h
32	17	13.8	117	23	AAO14981	Humanised murine h
33	17	13.8	119	22	AAE07034	Humanised murine a
34	17	13.8	250	23	ABP45333	Human Blys binding
35	17	13.8	494	22	AAW95696	Human protein sequ
36	16	13.0	114	13	AAW28748	Heavy chain variab
37	16	13.0	118	13	AAW25412	Heavy chain variab
38	16	13.0	120	14	AAW30763	Heavy chain variab
39	16	13.0	120	15	AAW47040	Heavy chain variab
40	16	13.0	121	12	AAAR3182	Sequence of the he
41	16	13.0	122	14	AAW30772	Variable region of
42	16	13.0	125	15	AAW56287	huxCD3v9, humanise
43	16	13.0	137	23	ABW77063	Variable heavy cha
44	16	13.0	239	23	ABP45871	AC2885 antibody he
45	16	13.0	249	23	ABP45806	Human Blys binding

#### ALIGNMENTS

RESULT 1	AAW01587	standard; Protein; 110 AA.
ID	AAW01587	
AC	AAW01587	
DT	22-AUG-1997	(first entry)
DE	Lead binding Mab 6F5 heavy chain variable region.	
XX	Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic; pharmaceutical; health care; skin treatment; pesticide; herbicide; heavy metal.	
XX		
OS	Mus musculus.	
XX		
PN	W09639518-A1.	
XX		
PD	12-DEC-1996.	
XX		
PF	05-JUN-1996;	96WO-US09258.
XX		
PR	10-OCT-1995;	95US-0541373.
XX		
PR	05-JUN-1995;	95US-0462798.
XX		
PA	(BION-) BIONEERASKA INC.	
XX		
PI	Lopez O, Murray PJ, Wylie DE;	
XX		
DR	WPI; 1997-043140/04.	
XX		
DR	N-PSDB; AAT58261.	
XX		
PT	DNA encoding heavy metal binding polypeptide sequences - used for	

PT detecting, removing, adding or neutralising heavy metals, such as  
 PT lead cations  
 XX  
 PS Claim 12: Page 77; 125pp; English.  
 XX  
 CC The present sequence represents the heavy chain variable region for  
 CC monoclonal antibody (MAb) 6F5, which immunoreacts with a lead cation.  
 CC The sequence was derived from RNA isolated from mouse hybridoma cells.  
 CC The protein can be used for binding heavy metals, such as lead cations.  
 CC It can be used for detecting, removing, adding or neutralising the  
 CC heavy metals in biological and inanimate systems. It can be used in  
 CC e.g. aqueous liquid systems, in biological or environmental systems or  
 CC in such compositions as perfumes, cosmetics, pharmaceuticals, health  
 CC care products, skin treatment products, pesticides, herbicides,  
 CC solvents used in the production of semi-conductor and integrated  
 CC circuit components and production materials for electronic components.  
 CC The products can provide for applications involving minute amounts of  
 CC specific heavy metals.  
 CC  
 SQ Sequence 110 AA;  
 XX  
 QY 62 YADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYYCVR 101  
 DB 57 YADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYYCVR 96  
 XX  
 RESULT 2  
 AAW01593  
 ID AAW01593 standard; Protein: 115 AA.  
 XX  
 AC AAW01593;  
 XX  
 DT 22-AUG-1997 (first entry)  
 XX  
 DE Lead binding MAb 2B4 heavy chain variable region.  
 XX  
 KM Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic;  
 KM pharmaceutical; health care; skin treatment; pesticide; herbicide;  
 KM heavy metal.  
 XX  
 OS Mus musculus.  
 XX  
 PN MO9639518-A1.  
 XX  
 PD 12-DEC-1996.  
 XX  
 PF 05-JUN-1996; 96WO-US09258.  
 XX  
 PR 10-OCT-1995; 95US-0541373.  
 PR 05-JUN-1995; 95US-0462798.  
 XX  
 PA (BION-) BIONEBRASKA INC.  
 XX  
 PI Lopez O, Murray PJ, Wylie DE;  
 XX  
 DR WPI: 1997-043140/04.  
 DR N-PSDB: AAT58267.  
 XX  
 XX DNA encoding heavy metal binding polypeptide sequences - used for  
 PT detecting, removing, adding or neutralising heavy metals, such as  
 PT lead cations  
 XX  
 PS Claim 12: Page 89; 125pp; English.  
 XX  
 CC The present sequence represents the heavy chain variable region for  
 CC monoclonal antibody (MAb) 2B4, which immunoreacts with a lead cation.  
 CC The sequence was derived from RNA isolated from mouse hybridoma cells.  
 CC The protein can be used for binding heavy metals, such as lead cations.  
 CC It can be used for detecting, removing, adding or neutralising the

CC heavy metals in biological and inanimate systems. It can be used in  
 CC e.g. aqueous liquid systems, in biological or environmental systems or  
 CC in such compositions as perfumes, cosmetics, pharmaceuticals, health  
 CC care products, skin treatment products, pesticides, herbicides,  
 CC solvents used in the production of semi-conductor and integrated  
 CC circuit components and production materials for electronic components.  
 CC The products can provide for applications involving minute amounts of  
 CC specific heavy metals.  
 CC  
 SQ Sequence 115 AA;  
 XX  
 QY 62 YADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYYCVR 101  
 DB 62 YADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYYCVR 101  
 XX  
 RESULT 3  
 AAE06972  
 ID AAE06972 standard; Protein: 100 AA.  
 XX  
 AC AAE06972;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Mouse germline heavy chain variable (VH) region, MRL-RF24B.  
 XX  
 KM Mouse: humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 KM neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KM HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KM inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
 KM multiple sclerosis; atherosclerosis; atherosclerosis; asthma;  
 KM anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KM fibrotic disease; angiodysplasia; acquired immune deficiency syndrome;  
 KM AIDS; inflammatory glomerulopathy; vascular intervention;  
 KM neointimal hyperplasia; VH: heavy chain variable region.  
 XX  
 OS Mus sp.  
 XX  
 PN MO200157226-A1.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 02-FEB-2001; 2001WO-US03537.  
 XX  
 PR 03-FEB-2000; 2000US-0497625.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 XX  
 DR WPI: 2001-48888/53.  
 XX  
 XX Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 PT CCR2, and a non-human antigen binding region and human immunoglobulin  
 PT -  
 XX  
 PS Disclosure, Page 152; 183pp; English.  
 XX  
 CC The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherosclerosis and atherosclerosis,

CC especially in gene therapy. The invention overcomes the restricted host

50 Sequence 32 AA:

Query Match	26.08;	Score 32;	DB 20;	Length 32;
-------------	--------	-----------	--------	------------

Best Local Similarity 100.0%; Pred. No. 2,6e-24;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 69 RFTISRDDSSMLYLQNNKTEDTAMYYCVR 100  
Db 1 RFTISRDDSSMLYLQNNKTEDTAMYYCVR 32

## RESULT 6

AAW94735  
ID AAW94735 standard; Protein; 30 AA.

AAW94735;

22-APR-1999 (first entry)

Anti-Staph (HAY) 96-110 heavy chain variable region.

Monoclonal antibody; MAb; lipoteichoic acid; gram positive; bacteria;  
immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
MAb 96-110.

Mus sp.

WO9857994-A2.

23-DEC-1998.

16-JUN-1998; 98WO-US12402.

16-JUN-1997; 97US-0049871.

(JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

Fischer GW, Schuman RF, Stinson JL, Wong H;

WPI; 1999-095329/08.

N-PSDB; AAX05579.

New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT infections caused by gram positive bacteria

Claim 21: Fig 12; 150pp; English.

The invention relates to a monoclonal antibody (MAb) to lipoteichoic  
acid of gram positive bacteria, where the MAb is a chimeric  
immunoglobulin comprising at least part of a human immunoglobulin  
constant region and at least part of a non-human immunoglobulin variable  
region having specificity to lipoteichoic acid of gram positive bacteria.  
The antibodies bind to whole bacteria and enhance phagocytosis and  
killing of the bacteria and enhance protection from lethal infection. The  
antibodies or peptides (encoded by a DNA of the variable region of  
anti-lipoteichoic acid antibody or characterised by amino acids  
corresponding to one or more of the complementarity determining  
regions (CDRs) of the variable region of the antibody) can be used for treating  
or preventing infections caused by gram positive bacteria. They can also  
be used for the diagnosis of gram positive bacterial infections.  
Sequences AAW94735-39 represent heavy chain variable regions of the  
anti-lipoteichoic antibody 96-100.

Sequence 30 AA;

Query Match 24.4%; Score 30; DB 20; Length 30;

Best Local Similarity 100.0%; Pred. No. 2,4e-22;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVMIVSGGGLVOPKSLKSCAASGFTFN 30  
Db 1 EVMIVSGGGLVOPKSLKSCAASGFTFN 30

## RESULT 7

AA90829  
ID AA90829 standard; Protein; 119 AA.

AA90829;

25-JUN-1996 (first entry)

3B1 heavy chain variable region from PCIB4613.

delta endotoxin; Bacillus thuringiensis; western corn rootworm;  
KCRW; maize; pesticide; brush border membrane vesicle; monoclonal;  
antibody.

Insecta sp.

WO9600783-A1.

11-JAN-1996.

20-JUN-1995; 95WO-IB00497.

28-JUN-1994; 94US-0267641.

(CIBA) CIBA GEIGY AG.

Carozzi NB, Koziel MG;

WPI; 1996-077494/08.

N-PSDB; AAT15725.

New monoclonal antibodies which bind insect gut proteins - used  
PT partic. with toxin moieties for the control of insect pests, partic.  
PT in plants

Claim 8; Page 50-51; 106pp; English.

AA90829-39 are monoclonal antibodies or a binding fragments produced by  
CC using insect guts, partic. insect brush border membranes (BBMs), esp.  
CC corn rootworm, as antigen; immunising a donor animal with the antigen;  
CC isolating immunocompetent B cells from the immunised animal; fusing B  
CC cells with a tumour cell line; isolating the fused cells, culturing them  
CC and cloning positive hybrid cells; and screening the hybrid cells for  
CC prodn. of the regulated Mabs. The Mabs, bind to the gut of a target insect  
CC but do not bind to mammalian BBMs. The DNA sequence can be operably  
CC linked to a toxin moiety, esp. selected from e.g. Bacillus toxins,  
CC Pseudomonas exotoxin and phytolectin, etc.. The Abs are useful for  
CC control of insect pests, e.g. Coleoptera, Diptera, Hymenoptera and  
CC Lepidoptera. The pesticidal compsn. is pref. applied to a plant, e.g.  
CC maize.

Sequence 119 AA;

Query Match 22.8%; Score 28; DB 17; Length 119;

Best Local Similarity 100.0%; Pred. No. 7,4e-20;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 AMNVMVROAPGKGLFWARIRKSNVAT 60  
Db 33 AMNVMVROAPGKGLFWARIRKSNVAT 60

## RESULT 8

AA90837  
ID AA90837 standard; Protein; 599 AA.

AA90837;

25-JUN-1996 (first entry)

3B1 single chain antibody from PCIB4631.

delta endotoxin; Bacillus thuringiensis; western corn rootworm;  
KCRW; maize; pesticide; brush border membrane vesicle; monoclonal;

KM antibody.  
 XX  
 OS Insecta sp.  
 PN WO9600783-A1.  
 XX  
 PD 11-JAN-1996.  
 XX  
 PE 20-JUN-1995; 95WO-1B00497.  
 XX  
 PR 28-JUN-1994; 94US-0267641.  
 XX  
 PA (CIBA ) CIBA GEIGY AG.  
 XX  
 PI Carozzi NB, Koziel MG;  
 XX  
 DR WPI; 1996-077494/08.  
 DR N-PSDB; AAT15733.  
 XX  
 PT New monoclonal antibodies which bind insect gut proteins - used  
 PT partic. with toxin moieties for the control of insect pests, partic.  
 PT in plants  
 PS Claim 8: Page 68-72; 106pp: English.  
 XX  
 CC AAP0829-39 are monoclonal antibodies or a binding fragments produced by  
 CC using insect guts, partic. insect brush border membranes (BBMs), esp.  
 CC corn rootworm, as antigen; immunising a donor animal with the antigen;  
 CC isolating immunocompetent B cells from the immunised animal; fusing B  
 CC cells with a tumour cell line; isolating the fused cells; culturing them  
 CC and cloning positive hybrid cells; and screening the hybrid cells for  
 CC prodn. of the required MAb. The MAb bind to the gut of a target insect  
 CC but do not bind to mammalian BBMs. The DNA sequence can be operably  
 CC linked to a toxin moiety, esp. selected from e.g. Bacillus toxins,  
 CC pseudomonas exotoxin and phytolaccin, etc.. The Abs are useful for  
 CC control of insect pests, e.g. Coleoptera, Diptera, Hymenoptera and  
 CC Lepidoptera. The pesticidal compsn. is pref. applied to a plant, e.g.  
 CC maize.  
 XX  
 SQ Sequence 599 AA:  
 Query Match 22.8%; Score 28; DB 17; Length 599;  
 Best Local Similarity 100.0%; Pred. No. 3e-19;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 AMNWVROARGKLEWVARIRSKSNNTAT 60  
 DB 187 AMNWVROARGKLEWVARIRSKSNNTAT 214  
 RESULT 9  
 ID AAE06971 standard; Protein: 101 AA.  
 AC AAE06971;  
 XX  
 DT 16-OCT-2001 (first entry)  
 DE Murine 1D9 antibody partial heavy chain variable (VH) region.  
 XX  
 KM Murine: humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 KM neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KM HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KM inflammatory disorder; autoimmune disorder; Rheumatoid arthritis; shock;  
 KM multiple sclerosis; atherosclerosis; stenosis; asthma;  
 KM anaphylaxis; malignancy; inflammation; sclerosis; allograft rejection;  
 KM fibrotic disease; angiodysplasia; acquired immune deficiency syndrome;  
 KM AIDS; inflammatory glomerulopathy; vascular intervention; 1D9 antibody;  
 KM neonatal hyperplasia; VH; heavy chain variable region.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200157226-A1.

XX  
 PD 09-AUG-2001.  
 XX  
 PF 02-FEB-2001; 2001WO-US03537.  
 XX  
 PR 03-FEB-2000; 2000US-0497625.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 XX  
 DR WPI; 2001-488888/53.  
 XX  
 PT Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 PT CCR2, and a non-human antigen binding region and human immunoglobulin  
 PT -  
 PS Disclosure: Page 152; 183pp: English.  
 XX  
 CC The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherosclerosis and atherosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neonatal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is partial heavy chain variable (VH) region of  
 CC murine 1D9 antibody.  
 XX  
 SQ Sequence 101 AA:  
 Query Match 19.5%; Score 24; DB 22; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-16;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 LVESGGGLVOPKSLKSCAASGF 27  
 DB 4 LVESGGGLVOPKSLKSCAASGF 27  
 RESULT 10  
 ID AAU09919 standard; Protein: 117 AA.  
 AC AAU09919;  
 XX  
 DT 18-JUN-2002 (first entry)  
 DE Murine mAb 1D9 heavy chain variable region protein sequence.  
 XX  
 KM Mouse; mAb 1D9 heavy chain variable region; vasotropic;  
 KM antinflammatory; collagen disease; immunosuppressive;  
 KM antidiabetic; insulin-dependent diabetes mellitus;  
 KM inflammatory bowel disease; ulcerative colitis;  
 KM graft rejection; allergic disease; antipsoriatic;  
 KM antiarrhythmic; nephrotropic; antichyroid; restenosis;  
 KM dermatological; anaphylaxis; cell adhesion inhibitor;  
 KM vascular injury; autoimmune disease; immunoglobulin;

KW		complementarity determining region; CDR; CD18; CCR2;
KW		atherosclerosis.
XX		
XX	Mus sp.	
OS		
FH	Key	Location/Qualifiers
FT	Region	/note= "Part of HL structure loop"
FT	Region	31..35
FT	Region	/note= "Complementarity determining region 1 (CDR1)"
FT	Region	50..68
FT	Region	/note= "Complementarity determining region 2 (CDR2)"
FT	Region	101..106
FT	Region	/note= "Complementarity determining region 3 (CDR3)"
PX		
PN	WO200170266-A2.	
PD	27-SEP-2001.	
PX		
PX	15-MAR-2001; 2001WO-US08266.	
PR	17-MAR-2000; 2000US-0528267.	
PX	(MILL-) MILLENNIUM PHARM INC.	
PA		
PX	Horvath CJ, Rao PE;	
PI		
PX	WPI; 2001-607511/69.	
DR		
PX		
PT	Inhibiting stenosis or restenosis of a blood vessel following vascular	
PT	injury or angioplasty in a subject by administering agent which	
PT	inhibits recruitment or adhesion of neutrophils, mononuclear cells to	
PT	injury site -	
PX		
PS	Claim 30; Fig 16; 108pp; English.	
XX		
XX	The present invention relates to a new method of inhibiting stenosis or	
CC	restenosis of a blood vessel following vascular injury in a subject. The	
CC	new method comprises administering to the subject agents which inhibit	
CC	the adhesion and/or recruitment of neutrophils and mononuclear cells to	
CC	a site of vascular injury by binding CD18 or CCR2. The method of the	
CC	invention inhibits stenosis or restenosis of a blood vessel following	
CC	vascular injury arising from a vascular intervention procedure such as	
CC	vascular by-pass or transplantation surgery. The method is also useful	
CC	for treating a subject having an inflammatory disease or condition	
CC	mediated by neutrophil and mononuclear cell activity e.g. asthma and	
CC	graft versus host disease. Chronic inflammatory diseases of the lung,	
CC	collagen diseases, and insulin-dependent diabetes mellitus can also be	
CC	treated. The method is further useful for treating inflammatory bowel	
CC	diseases, such as ulcerative colitis. Additional diseases or conditions	
CC	include inflammation or allergic diseases and conditions, including	
CC	systemic anaphylaxis of hypersensitivity responses, drug allergies,	
CC	psoriasis and inflammatory dermatoses, autoimmune diseases such as	
CC	arthritis, graft rejection and other diseases including atherosclerosis.	
CC	The present sequence represents the protein sequence of the mouse mab	
CC	1D9 heavy chain variable region. This sequence contains three CDR's	
CC	that can be used to inhibit stenosis or restenosis.	
CC		
SQ	Sequence 117 AA:	
OY	Query Match 19.5%; Score 24; DB 22; Length 117;	
Db	Best Local Similarity 100.0%; Pred. No. 6,66-16;	
	Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
	4 LVESGGGLVOPKGSILKLSCAASGF 27	
	4 LVESGGGLVOPKGSILKLSCAASGF 27	
RESULT 11		
AABE0947		
ID AAE06947 standard; Protein; 117 AA.		
XX		

AC	AAE06947;
XX	
XX	16-OCT-2001 (first entry)
DE	
XX	Murine ID9 antibody heavy chain variable (VH) region.
XX	
KM	Murine; humanised antibody; CC-chemokine receptor 2; CCR2; neutrotropic
KM	neuroprotective; immunosuppressive; human immunodeficiency virus;
KM	HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;
KM	inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock
KM	multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;
KM	anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;
KM	fibrotic disease; angiodysplasia; acquired immune deficiency syndrome;
KM	AIDS; inflammatory glomerulopathy; vascular intervention; ID9 antibody;
KM	neointimal hyperplasia; VH; heavy chain variable region.
XX	
OS	Mus sp.
XX	
FH	Key
FT	Region
FT	31..35
FT	/label- Complementarity_determining_region
FT	50..68
FT	/label- Complementarity_determining_region
FT	101..106
FT	/label- Complementarity_determining_region
XX	
PN	WO200157226-A1.
XX	
PD	09-AUG-2001.
XX	
PF	02-FEB-2001; 2001WO-US03537.
XX	
PF	03-FEB-2000; 2000US-0497625.
PR	
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;
XX	WPI: 2001-488888/53.
DR	
XX	
PT	Humanized immunoglobulin for treating a CC-chemokine receptor
PT	2-mediated disorder in a patient, comprises a binding specificity for
PT	CCR2, and a non-human antigen binding region and human immunoglobulin
PT	-
PS	Claim 8; Fig 8; 183pp; English.
XX	
CC	The patent discloses a humanised antibody or its antigen-binding
CC	fragment, having binding specificity for CC-chemokine receptor 2
CC	(CCR2), comprising an antigen binding region of non-human origin
CC	and at least a portion of an immunoglobulin of human origin. The
CC	humanised antibodies are useful for inhibiting the interaction of
CC	a cell expressing CCR2. They are useful for inhibiting or creating
CC	Leukocyte trafficking, for treating CCR2-mediated disorders such as
CC	inflammatory disorder, autoimmune disorders such as rheumatoid
CC	arthritis and multiple sclerosis, atherogenesis and atherosclerosis,
CC	and for inhibiting restenosis. They are useful in therapy or diagnosis,
CC	and in the manufacture of a medicament for treating CCR2 mediated
CC	disease. They are also useful for treating allergy, anaphylaxis,
CC	maligancy, chronic and acute inflammation, histamine and IgE-
CC	mediated allergic reaction, shock, stenosis, allograft rejection,
CC	fibrotic disease, asthma, inflammatory glomerulopathies, acquired
CC	immune deficiency syndrome (AIDS), restenosis associated with vascular
CC	intervention, including angioplasty and/or stent placement in a mammal.
CC	Humanised antibodies are also useful for inhibiting narrowing of the
CC	lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of
CC	a vessel in a mammal, preferably associated with vascular intervention.
CC	The present sequence is the heavy chain variable (VH) region of
CC	murine ID9 antibody.
XX	
50	Sequence 117 AA;

Query Match 19.5%; Score 24; DB 22; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-16;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LVESGGGLVQPKGSLKLSCAASGF 27  
 |||  
 DB 4 LVESGGGLVQPKGSLKLSCAASGF 27

RESULT 12  
 AA014978  
 ID AA014978 standard; protein: 117 AA.  
 XX  
 AC AA014978;  
 XX  
 DT 05-SEP-2002 (first entry)  
 DE Mouse heavy chain variable region (mab 1D9).  
 XX  
 KW Mouse; graft rejection; CC chemokine receptor 2 antagonist;  
 KW CCR2 antagonist; anti-CCR2 antibody; kidney transplant; liver transplant;  
 KW lung transplant; heart-lung transplant; pancreas transplant;  
 KW bowel transplant; heart transplant; graft versus host disease;  
 KW chronic graft rejection; antibody heavy chain variable region; mab 1D9.  
 OS Mus musculus.  
 XX  
 PM US2002042370-A1.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 13-APR-2001; 2001US-0835087.  
 XX  
 PR 14-APR-2000; 2000US-0549448.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Hancock MW;  
 XX  
 DR WPI: 2002-351265/38.  
 XX  
 PT Inhibiting graft rejection, graft versus host disease or chronic  
 PT rejection of a transplanted graft, involves administering a CCR2  
 PT antagonist -  
 XX  
 PS Disclosure: Fig 2; 16pp; English.  
 XX  
 CC The invention comprises a method of inhibiting graft rejection, graft  
 CC versus host disease or chronic rejection of a transplanted graft. The  
 CC method involves administering an antagonist of CC chemokine receptor 2  
 CC (CCR2) and optionally an immunosuppressive agent. The CCR2 antagonist may  
 CC be an anti-CCR2 antibody (i.e. containing light and heavy chain  
 CC complementarily determining regions from various non-human origins). CCR2  
 CC is known to be involved in the rejection of transplanted grafts. The  
 CC method of the invention is useful for inhibiting graft rejection -  
 CC particularly allografts such as kidney, liver, lung, heart-lung,  
 CC pancreas, bowel and heart. The method of the invention is also useful for  
 CC inhibiting graft versus host disease and for inhibiting chronic rejection  
 CC of a transplanted graft. The present amino acid sequence represents a  
 CC mouse antibody heavy chain variable region (mab 1D9).  
 XX  
 SQ Sequence 117 AA;  
 Query Match 19.5%; Score 24; DB 23; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-16;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LVESGGGLVQPKGSLKLSCAASGF 27  
 |||  
 DB 4 LVESGGGLVQPKGSLKLSCAASGF 27

RESULT 13

AAE07033  
 ID AAE07033 standard; Protein: 148 AA.  
 XX  
 AC AAE07033;  
 XX  
 DT 16-OCT-2001 (first entry)  
 DE Murine antibody 1D9 heavy chain variable region.  
 XX  
 KW Murine; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
 KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention;  
 KW neointimal hyperplasia; antibody 1D9 heavy chain variable region.  
 XX  
 OS Mus sp.  
 XX  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FH Peptide 1..19  
 FT /Label= Signal\_peptide  
 FT Protein 20..148  
 FT /note= "Murine mature antibody 1D9 kappa  
 FT light chain variable region"  
 FT Region 20..137  
 FT /Label= Variable\_region  
 FT Region 138..148  
 FT /Label= constant\_region  
 FT Misc-difference 148  
 FT /note= "Encoded by Gt"  
 XX  
 PN WO200157226-A1.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 02-FEB-2001; 2001WO-US03537.  
 XX  
 PR 03-FEB-2000; 2000US-0497625.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 DR WPI: 2001-488888/53.  
 DR N-PSDB: AAD13178.  
 XX  
 PT Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 PT CCR2, and a non-human antigen binding region and human immunoglobulin  
 XX  
 PS Disclosure: Fig 21; 183pp; English.  
 XX  
 CC The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular

CC Intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neonatal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is murine antibody 1D9 heavy chain variable region.  
 XX  
 SQ Sequence 148 AA;  
 Query Match 19.5%; Score 24; DB 22; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 8, 1e-16;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 4 LVESGGGLVOPKSGSLKSCASGF 27  
 Db 23 LVESGGGLVOPKSGSLKSCASGF 46  
 RESULT 14  
 ID AAW94739 standard; Protein; 23 AA.  
 XX AAW94739:  
 AC AAW94739:  
 XX  
 DT 22-APR-1999 (first entry)  
 DE Anti-Staph (HAY) 96-110 heavy chain variable region.  
 XX  
 KM Monoclonal antibody; Mab: lipoteichoic acid; gram positive; bacteria;  
 KM immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KM MAb 96-110.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..12  
 FT /note= "complementarity determining region (CDR)"  
 XX  
 PN WO9857994-A2.  
 PD 23-DEC-1998.  
 XX  
 PF 16-JUN-1998; 98WO-US12402.  
 XX  
 PR 16-JUN-1997; 97US-0049871.  
 XX  
 PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
 XX  
 PI Fischer GW, Schuman RF, Stinson JL, Wong H;  
 DR WPI; 1999-095329/08.  
 DR N-PSDB; AAX05583.  
 XX  
 PT New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 XX  
 PS Claim 21; Fig 12; 150pp; English.  
 XX  
 CC The invention relates to a monoclonal antibody (MAB) to lipoteichoic  
 CC acid of gram positive bacteria, where the MAB is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94735-39 represent heavy chain variable regions of the  
 CC anti-lipoteichoic antibody 96-100.

XX  
 SQ Sequence 23 AA;  
 Query Match 18.7%; Score 23; DB 20; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 1, 6e-15;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 101 RGASGIDYAMDYWGQSTLTYS 123  
 Db 1 RGASGIDYAMDYWGQSTLTYS 23  
 RESULT 15  
 ID AAU09928 standard; Protein; 117 AA.  
 XX AAU09928:  
 AC AAU09928:  
 XX  
 DT 18-JUN-2002 (first entry)  
 DE Humanised 1D9 heavy chain variable region, 1D9RHB protein sequence.  
 XX  
 KM Human; mouse; 4B4/CL heavy chain variable region; vasotropic;  
 KM antiinflammatory; collagen disease; immunosuppressive;  
 KM antidiabetic; insulin-dependent diabetes mellitus;  
 KM inflammatory bowel disease; ulcerative colitis; 1D9RHB;  
 KM graft rejection; allergic disease; antipsoriatic;  
 KM arthritic; nephrotropic; cell adhesion inhibitor;  
 KM dermatological; anaphylaxis; cell adhesion inhibitor;  
 KM vascular injury; autoimmune disease; immunoglobulin;  
 KM complementarity determining region; CDR; CD18; CCR2;  
 KM atherosclerosis; mutant; mutein.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Mus sp.  
 OS Chimeric - Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 27..30  
 FT /note= "Part of H1 structure loop"  
 XX  
 FT Misc-difference 28  
 FT /note= "Substitution of Thr residue normally present  
 FT in human 4B4/CL sequence (AAU09926) by Ser  
 FT residue normally present in mouse mab 1D9 heavy  
 FT chain sequence (AAU09919)"  
 XX  
 FT Misc-difference 30  
 FT /note= "Substitution of Ser residue normally present  
 FT in human 4B4/CL sequence (AAU09926) by Asn  
 FT residue normally present in mouse mab 1D9 heavy  
 FT chain sequence (AAU09919)"  
 XX  
 FT Region 31..35  
 FT /note= "Complementarity determining region 1 (CDR1),  
 FT grafted from mouse mab 1D9 heavy chain  
 FT sequence (AAU09919)"  
 XX  
 FT Region 50..68  
 FT /note= "Complementarity determining region 2 (CDR2),  
 FT grafted from mouse mab 1D9 heavy chain  
 FT sequence (AAU09919)"  
 XX  
 FT Region 101..108  
 FT /note= "Complementarity determining region 3 (CDR3),  
 FT grafted from mouse mab 1D9 heavy chain  
 FT sequence (AAU09919)"  
 XX  
 PN WO200170266-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PR 15-MAR-2001; 2001WO-US08266.  
 XX  
 PR 17-MAR-2000; 2000US-0528267.  
 XX



PA (MILL-) MILLENNIUM PHARM INC.

XX  
XX  
PI Horvath CJ, Rao PE;

XX  
XX  
DR WPI; 2001-607511/69.

XX  
XX  
PT Inhibiting stenosis or restenosis of a blood vessel following vascular  
PT injury or angioplasty in a subject by administering agent which  
PT inhibits recruitment or adhesion of neutrophils, mononuclear cells to  
PT injury site -

PS Claim 32; Fig 18; 108pp; English.

XX  
XX  
CC The present invention relates to a new method of inhibiting stenosis or  
CC restenosis of a blood vessel following vascular injury in a subject. The  
CC new method comprises administering to the subject agents which inhibit  
CC the adhesion and/or recruitment of neutrophils and mononuclear cells to  
CC a site of vascular injury by binding CD18 or CCR2. The method of the  
CC invention inhibits stenosis or restenosis of a blood vessel following  
CC vascular injury arising from a vascular intervention procedure such as  
CC vascular by-pass or transplantation surgery. The method is also useful  
CC for treating a subject having an inflammatory disease or condition  
CC mediated by neutrophil and mononuclear cell activity e.g. asthma and  
CC graft versus host disease. Chronic inflammatory diseases of the lung,  
CC collagen diseases, and insulin-dependent diabetes mellitus can also be  
CC treated. The method is further useful for treating inflammatory bowel  
CC diseases, such as ulcerative colitis. Additional diseases or conditions  
CC include inflammatory or allergic diseases and conditions, including  
CC systemic anaphylaxis of hypersensitivity responses, drug allergies,  
CC psoriasis and inflammatory dermatoses, autoimmune diseases such as  
CC arthritis, graft rejection and other diseases including atherosclerosis.  
CC This sequence represents the variable region of one of several humanised  
CC 1D9 heavy chains (AAU09927-AAU09930). These heavy chains were used in  
CC the invention for the production of anti-CCR2 antibody or antigen-binding  
CC fragment.

XX  
SQ Sequence 117 AA;

Query Match 17.1%; Score 21; DB 22; Length 117;

Best Local Similarity 100.0%; Pred. No. 6.1e-13;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 YAMNWYROAPGKGLEWVARIR 52

DB 32 YAMNWYROAPGKGLEWVARIR 52

Search completed: November 27, 2002, 07:27:59  
Job time : 45.6911 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OK protein - protein search, using sw model

Run on: November 27, 2002, 07:25:20 ; Search time 18.5212 Seconds  
(without alignments)  
638.431 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 123

Sequence: 1 EVMLESGLVGPCKSLK.....SGIDYAMDYMGQTSLEFVSS 123

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	31.7	119	2	A27630 Ig heavy chain pre
2	39	31.7	141	2	I32513 Ig heavy chain pre
3	38	30.9	100	2	S26462 Ig heavy chain v r
4	37	30.1	110	2	PH1091 Ig heavy chain v r
5	35	28.5	137	2	S42467 Ig heavy chain v r
6	28	22.8	110	2	PH1092 Ig heavy chain v r
7	28	22.8	121	2	A41940 Ig heavy chain v r
8	25	20.3	101	2	B42575 Ig heavy chain v r
9	20	16.3	110	2	PH1093 Ig heavy chain v r
10	18	14.6	118	2	S00700 Ig heavy chain v r
11	17	13.8	120	2	E49590 Ig heavy chain v r
12	16	13.0	105	2	S38488 Ig heavy chain v r
13	16	13.0	139	2	S31678 Ig heavy chain v r
14	15	12.2	34	2	S46469 Ig heavy chain v r
15	15	12.2	95	2	S20777 Ig heavy chain v r
16	15	12.2	98	2	PH0874 Ig heavy chain v r
17	15	12.2	98	2	S26891 Ig heavy chain v r
18	15	12.2	98	2	S26940 Ig heavy chain v r
19	15	12.2	98	2	S26894 Ig heavy chain v r
20	15	12.2	110	2	S36282 Ig heavy chain v r
21	15	12.2	112	2	S25572 Ig heavy chain v r
22	15	12.2	113	2	S25573 Ig heavy chain v r
23	15	12.2	116	2	S21979 Ig heavy chain v r
24	15	12.2	116	2	S17080 Ig heavy chain v r
25	15	12.2	117	2	S34012 Ig heavy chain v r
26	15	12.2	117	2	S21980 Ig heavy chain v r
27	15	12.2	119	2	PH1290 Ig heavy chain v r
28	15	12.2	119	2	PH1292 Ig heavy chain pre
29	15	12.2	121	2	H36005 Ig heavy chain v r

30	15	12.2	123	2	S26794 Ig heavy chain v r
31	15	12.2	123	2	S34009 Ig heavy chain v r
32	15	12.2	127	2	PH0369 Ig gamma chain pre
33	15	12.2	127	2	S19878 Ig heavy chain v r
34	15	12.2	133	2	PH1286 Ig heavy chain pre
35	15	12.2	141	2	PH1287 Ig heavy chain pre
36	15	12.2	143	2	S31669 Ig heavy chain v r
37	15	12.2	143	2	S23624 Ig heavy chain v r
38	15	12.2	144	2	PH1285 Ig heavy chain pre
39	15	12.2	145	2	I47203 Ig heavy chain pre
40	15	11.4	40	2	S33406 Ig heavy chain var
41	14	11.4	40	2	S33410 Ig heavy chain v r
42	14	11.4	83	2	C25913 Ig heavy chain v r
43	14	11.4	94	2	PH0120 Ig heavy chain v r
44	14	11.4	94	2	S14580 Ig heavy chain v r
45	14	11.4	96	2	PH0873 Ig heavy chain v r

## ALIGNMENTS

RESULT 1  
A27630

Ig heavy chain precursor V-10 region (MRL-DNA4) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 23-Jul-1999

C:Accession: A27630

R:Kotler, R.

J. Immunol. 140, 4031-4034, 1988

A:Title: A new murine Ig V-H gene family.

A:Reference number: A27630; MUID:88229103; PMID:3131427

A:Accession: A27630

A:Molecule type: DNA

A:Residues: 1-119 <KOR>

A:Cross-references: GB:M21470; NID:g196220; PIDN:AAA38634.1; PID:g196221; GB:M20825

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotrimer; Immunoglobulin

F:34-119/Domain: Immunoglobulin homology <IMM>

Query Match 31.7%; Score 39; DB 2; Length 119;

Best Local Similarity 100.0%; Pred. No. 3.5e-32;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 YADSVKDRFTISRDSQSMLYLQNMNLTREPTAMYYCVR 100

DB 81 YADSVKDRFTISRDSQSMLYLQNMNLTREPTAMYYCVR 119

RESULT 2  
I32513

Ig heavy chain precursor V region (MRL4) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-May-1990 #sequence\_revision 31-Dec-1990 #text\_change 23-Jul-1999

C:Accession: I32513

R:Kotler, R.; Strobel, R.; Balderas, R.S.; Johnson, M.E.; Noonan, D.J.; Duchosal, M.A.

J. Clin. Invest. 82, 852-860, 1988

A:Title: Immunoglobulin kappa light chain variable region gene complex organization a

A:Reference number: A94689; MUID:8831394; PMID:3136286

A:Accession: I32513

A:Molecule type: DNA

A:Residues: 1-141 <KOR>

A:Cross-references: GB:M20829; NID:g196951; PIDN:AAA38649.1; PID:g196952

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotrimer; Immunoglobulin

F:34-119/Domain: Immunoglobulin homology <IMM>

Query Match 31.7%; Score 39; DB 2; Length 141;

Best Local Similarity 100.0%; Pred. No. 4e-32;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 YADSVKDRFTISRDSQSMLYLQNMNLTREPTAMYYCVR 100

DB 81 YADSVKDRFTISRDSQSMLYLQNMNLTREPTAMYYCVR 119

RESULT 3  
S26462  
Ig heavy chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
C:Accession: S26462  
R:Kavaler, J.  
submitted to the EMBL Data Library, April 1991  
A:Reference number: S26459  
A:Accession: S26462  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-100 <NAV>  
A:Cross-references: EMBL:X59104; NID:951920; PID:951921  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:7-92/Domain: Immunoglobulin homology <IMM>

Query Match 30.9%; Score 38; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.1e-31;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 YADSVKDRFTISRDSQSMYLQNNLKTEDTAMVYCV 99  
Db 54 YADSVKDRFTISRDSQSMYLQNNLKTEDTAMVYCV 91

RESULT 4  
PH1091  
Ig heavy chain V region (clone p17.79) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PH1091  
R:Fillman, D.M.; Jou, N.T.; Hill, R.J.; Marlon, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B  
A:Reference number: PH0971; MUID:92381444; PMID:1512540  
A:Accession: PH1091  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-110 <RTL>  
A:Experimental source: B cell, strain [NZB x NZW]F1  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match 30.1%; Score 37; DB 2; Length 110;  
Best Local Similarity 100.0%; Pred. No. 3.5e-30;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 YADSVKDRFTISRDSQSMYLQNNLKTEDTAMVYCV 98  
Db 62 YADSVKDRFTISRDSQSMYLQNNLKTEDTAMVYCV 98

RESULT 5  
S42467  
Ig heavy chain V region precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S42467  
R:Shlyanov, P.A.; Bspalov, I.A.; Terletska, H.N.; Deyev, S.M.  
submitted to the EMBL Data Library, March 1994  
A:Reference number: S42466  
A:Accession: S42467  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-137 <SHI>  
A:Cross-references: EMBL:X78107; NID:9460798; PID:CA54997.1; PID:9460799  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

F:34-119/Domain: Immunoglobulin homology <IMM>

Query Match 28.5%; Score 35; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 4.6e-28;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 SYKDRFTISRDSQSMYLQNNLKTEDTAMVYCV 99  
Db 84 SYKDRFTISRDSQSMYLQNNLKTEDTAMVYCV 118

RESULT 6  
PH1092  
Ig heavy chain V region (clone p17.80) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PH1092  
R:Fillman, D.M.; Jou, N.T.; Hill, R.J.; Marlon, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective  
A:Reference number: PH0971; MUID:92381444; PMID:1512540  
A:Accession: PH1092  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-110 <RTL>  
A:Experimental source: B cell, strain [NZB x NZW]F1  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match 22.8%; Score 28; DB 2; Length 110;  
Best Local Similarity 100.0%; Pred. No. 5.2e-21;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 YADSVKDRFTISRDSQSMYLQNNLK 89  
Db 62 YADSVKDRFTISRDSQSMYLQNNLK 89

RESULT 7  
A41940  
Ig heavy chain V region G2D, autoantibody BV04-01 - mouse (fragment)  
N:Alternate names: anti-DNA autoantibody BV04-01, heavy chain V region  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-May-1999  
C:Accession: A41940; PL0201  
R:Heiron, J.N.; He, X.W.; Ballard, D.W.; Blier, P.R.; Pace, P.E.; Bothwell, A.L.; Vos  
Protein 11, 159-175, 1991  
A:Title: An autoantibody to single-stranded DNA: comparison of the three-dimensional  
A:Reference number: A41940; MUID:92086633; PMID:1749770  
A:Accession: A41940  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-121 <HER>  
A:Note: sequence extracted from NCBI backbone (NCBI:P:70715)  
R:Smith, R.G.; Voss Jr., E.W.  
Mol. Immunol. 27, 463-470, 1990  
A:Title: Variable region primary structures of monoclonal anti-DNA autoantibodies fro  
A:Reference number: PL0198; MUID:90309768; PMID:2114528  
A:Accession: PL0201  
A:Molecule type: mRNA  
A:Residues: 1-118 <SMI>  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-100/Domain: Immunoglobulin homology <IMM>  
F:31-35/Region: complementarity-determining 1  
F:50-68/Region: complementarity-determining 2  
F:101-110/Region: complementarity-determining 3  
F:101-106/Region: D region  
F:107-115/Region: JH region

Query Match 22.8%; Score 28; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 5.6e-21;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 AMNVRQAPGKGLVAVIRSRKSNMYAT 60  
|||||  
Db 33 AMNVRQAPGKGLVAVIRSRKSNMYAT 60

RESULT 8  
B42575  
Ig heavy chain V region (anti-angiotensin II) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 21-Jan-2000  
C:Accession: B42575  
R:Garcla, K.C.; Desiderio, S.V.; Ronco, P.M.; Verroust, P.J.; Amzel, L.M.  
Science 257, 528-531, 1992  
A:Title: Recognition of angiotensin II: antibodies at different levels of an idiotypic T  
A:Reference number: A42575; MUID:92342952; PMID:1636087  
A:Accession: B42575  
A:Molecule type: mRNA  
A:Residues: 1-101 <GAR>  
A:Cross-references: GB:5A0679; NID:9252098; PIDN:AAB22669.1; PID:9252099  
A:Experimental source: BALB/c mice, myeloma cell line NS-1  
A:Note: sequence extracted from NCBI backbone (NCBI:109337, NCBI:P:109338)  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: Immunoglobulin  
F:7-92/Domain: Immunoglobulin homology <IMM>

Query Match 20.3%; Score 25; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 5.5e-18;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 YADSVKDRFTISRDSQSMYLOM 86  
|||||  
Db 54 YADSVKDRFTISRDSQSMYLOM 78

RESULT 9  
PH1093  
Ig heavy chain V region (clone S17.161) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PH1093  
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marlon, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B c  
A:Reference number: PH0971; MUID:92381444; PMID:1512540  
A:Accession: PH1093  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-110 <TIL>  
A:Experimental source: B cell, strain [NZB x NZW]F1  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin  
F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match 16.3%; Score 20; DB 2; Length 110;  
Best Local Similarity 100.0%; Pred. No. 7.3e-13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 GGGVQPKGSLKSCASGF 27  
|||||  
Db 8 GGGVQPKGSLKSCASGF 27

RESULT 10  
S00700  
Ig heavy chain V region - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 23-Jul-1999  
C:Accession: S00700  
R:Yamawaki-Kataoka, Y.; Honjo, T.  
Nucleic Acids Res. 15, 5888, 1987  
A:Title: Nucleotide sequences of variable region segments of the immunoglobulin heavy ch

A:Reference number: S00700; MUID:87289054; PMID:3112743  
A:Accession: S00700  
A:Molecule type: DNA  
A:Residues: 1-118 <YAM>  
A:Cross-references: EMBL:Y00380; NID:964810; PIDN:CAA68452.1; PID:964811  
A:Note: the sequence was determined from the germline gene  
C:Genetics:  
A:Introns: 15/3  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin  
F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 14.6%; Score 18; DB 2; Length 118;  
Best Local Similarity 100.0%; Pred. No. 8.4e-11;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 LYLQNNLKTEDTAMVYC 98  
|||||  
Db 98 LYLQNNLKTEDTAMVYC 115

RESULT 11  
E49590  
Ig heavy chain V region (ACHSV1, clone 2) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 16-Aug-1996  
C:Accession: E49590  
R:Burton, R.; Williamson, R.A.; Sanna, P.P.; Bloom, F.E.; Burton, D.R.  
Proc. Natl. Acad. Sci. U.S.A. 91, 355-359, 1994  
A:Title: Recombinant human Fab to glycoprotein D neutralizes infectivity and prevents  
A:Reference number: A49590; MUID:94105168; PMID:8278393  
A:Accession: E49590  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-120 <BUR>  
A:Experimental source: bone marrow lymphocytes  
A:Note: sequence extracted from NCBI backbone (NCBI:P:141854)  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin  
F:11-94/Domain: Immunoglobulin homology <IMM>

Query Match 13.8%; Score 17; DB 2; Length 120;  
Best Local Similarity 100.0%; Pred. No. 8.9e-10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 YAMNVRQAPGKGLV 48  
|||||  
Db 28 YAMNVRQAPGKGLV 44

RESULT 12  
S38488  
Ig heavy chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S38488  
R:Marks, J.D.; Ouehand, W.H.; Bye, J.M.; Finern, R.; Corick, B.D.; Voak, D.; Thorpe  
submitted to the EMBL Data Library, June 1993  
A:Description: Human antibody fragments specific for human blood group antigens from  
A:Reference number: S38488  
A:Accession: S38488  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-105 <MAR>  
A:Cross-references: EMBL:Z23026; NID:9414023; PIDN:CAA64561.1; PID:9414024  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin  
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 13.0%; Score 16; DB 2; Length 105;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 MNWVROAPGKLEWV 48  
 |||||||||||||  
 DB 33 MNWVROAPGKLEWV 48

## RESULT 13

S31678

Ig heavy chain V region - human

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999

C:Accession: S31678

R:Cuisinier, A.M.; Gauthier, L.; Boudli, L.; Fougereau, M.; Tonnelie, C.

submitted to the EMBL Data Library, June 1992

A:Description: Mechanisms that generate human immunoglobulin diversity operate from the

A:Reference number: S31585

A:Accession: S31678

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-139 &lt;CUI&gt;

A:Cross-references: EMBL:214169; NID:g31003; PIDN:CAA78538.1; PID:g31004

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin

F:31-116/Domain: Immunoglobulin homology &lt;IMM&gt;

Query Match 13.0%; Score 16; DB 2; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 1e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WVRQAPGKLEWVARI 51  
 |||||||||||||  
 DB 52 WVRQAPGKLEWVARI 67

## RESULT 14

S46469

Ig heavy chain V region (DA-5) - human

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000

C:Accession: S46469

R:COOK, G.P.; Tomlinson, I.M.; Walter, G.; Rietman, H.; Carter, N.P.; Bulwela, L.; Win

Nature Genet. 7, 162-168, 1994

A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the telom

A:Reference number: S46460; MUID:95004581; PMID:7920635

A:Accession: S46469

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-34 &lt;COO&gt;

A:Cross-references: EMBL:229981; NID:g505463; PIDN:CAA82871.1; PID:g1335180

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin

Query Match 12.2%; Score 15; DB 2; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-08;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 MNWVROAPGKLEWV 48  
 |||||||||||||  
 DB 7 MNWVROAPGKLEWV 21

## RESULT 15

S20777

Ig heavy chain V region - human

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999

C:Accession: S20777

R:Mortari, F.; Wang, J.; Schroeder, H.W.

submitted to the EMBL Data Library, April 1992

A:Description: Analysis of human cord blood Ig heavy chain Iga and Igg repertoire.

A:Reference number: S20764

A:Accession: S20777

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-95 <MOR>  
 A:Cross-references: EMBL:211947; NID:g33886; PIDN:CAA78004.1; PID:g33887  
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
 C:Keywords: heterotetramer; Immunoglobulin

Query Match 12.2%; Score 15; DB 2; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 MNWVROAPGKLEWV 48  
 |||||||||||||  
 DB 9 MNWVROAPGKLEWV 23

Search completed: November 27, 2002, 07:31:21  
 Job time : 19.5212 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:23:05 ; Search time 11.8726 Seconds  
(without alignments)  
429.695 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 123

Sequence: 1 EVMLVESGGGLVQPKGSLK.....SGIDYAMDYWGCGTSLTVSS 123

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SWISSProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	11.4	97	1 HV56_MOUSE	P18527 mus musculu
2	14	11.4	98	1 HV57_MOUSE	P18528 mus musculu
3	14	11.4	114	1 HV3B_HUMAN	P01763 homo sapien
4	14	11.4	116	1 HV3T_HUMAN	P01781 homo sapien
5	14	11.4	117	1 HV3O_HUMAN	P01776 homo sapien
6	14	11.4	117	1 HV53_MOUSE	P18524 mus musculu
7	14	11.4	117	1 HV54_MOUSE	P18525 mus musculu
8	14	11.4	119	1 HV3I_HUMAN	P01770 homo sapien
9	14	11.4	119	1 HV3M_HUMAN	P01774 homo sapien
10	14	11.4	119	1 HV3N_HUMAN	P01775 homo sapien
11	14	11.4	121	1 HV3J_HUMAN	P01771 homo sapien
12	14	11.4	126	1 HV3K_HUMAN	P01772 homo sapien
13	14	11.4	142	1 HV01_RAT	P01605 rattus norv
14	13	10.6	115	1 HV3S_HUMAN	P01780 homo sapien
15	13	10.6	116	1 HV3Q_HUMAN	P01778 homo sapien
16	13	10.6	116	1 HV3R_HUMAN	P01779 homo sapien
17	13	10.6	117	1 HV03_CARAU	P19180 carassius a
18	13	10.6	117	1 HV3C_HUMAN	P01764 homo sapien
19	13	10.6	118	1 HV3Y_HUMAN	P80419 homo sapien
20	13	10.6	119	1 HV3P_HUMAN	P01760 homo sapien
21	13	10.6	124	1 HV1D_HUMAN	P01760 homo sapien
22	13	10.6	136	1 HV2C_RABIT	P01829 oryctolagus
23	13	10.6	136	1 HV1A_RABIT	P01826 oryctolagus
24	12	9.8	116	1 HV36_MOUSE	P01806 mus musculu
25	12	9.8	117	1 HV2B_RABIT	P01828 oryctolagus
26	12	9.8	117	1 HV41_MOUSE	P01811 mus musculu
27	12	9.8	117	1 HV42_MOUSE	P01812 mus musculu
28	12	9.8	117	1 HV55_MOUSE	P18526 mus musculu
29	12	9.8	119	1 HV37_MOUSE	P01807 mus musculu
30	12	9.8	119	1 HV38_MOUSE	P01808 mus musculu
31	12	9.8	119	1 HV40_MOUSE	P01810 mus musculu
32	11	8.9	116	1 HV01_HETFR	P03983 heterodontu
33	11	8.9	117	1 HV17_MOUSE	P01786 mus musculu

## ALIGNMENTS

34	11	8.9	120	1 HV3E_HUMAN	P01766 homo sapien
35	11	8.9	122	1 HV20_MOUSE	P01789 mus musculu
36	11	8.9	122	1 HV21_MOUSE	P01790 mus musculu
37	11	8.9	123	1 HV18_MOUSE	P01787 mus musculu
38	11	8.9	123	1 HV19_MOUSE	P01788 mus musculu
39	11	8.9	123	1 HV22_MOUSE	P01791 mus musculu
40	11	8.9	123	1 HV23_MOUSE	P01792 mus musculu
41	11	8.9	123	1 HV24_MOUSE	P01793 mus musculu
42	11	8.9	123	1 HV25_MOUSE	P01794 mus musculu
43	11	8.9	136	1 HV16_MOUSE	P01783 mus musculu
44	10	8.1	114	1 HV01_CANFA	P01783 canis fam11
45	10	8.1	115	1 HV3D_HUMAN	P01765 homo sapien

## RESULT 1

AC	P18527;	STANDARD;	PRT;	97 AA.
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig heavy chain V region 914.			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/CJ;			
RX	MEDLINE=89279149; PubMed=2499654;			
RA	Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;			
RT	"Early onset of somatic mutation in immunoglobulin VH genes during			
RT	the primary immune response."			
RL	J. Exp. Med. 169:2007-2019(1989).			
CC	-I- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.			
DR	PIR: J0504; HYMS91.			
DR	HSSP: P01810; 2FB3.			
DR	InterPro: IPR003006; Ig_MHC.			
DR	InterPro: IPR003596; Ig_V.			
DR	Pfam: PF00047; 1g_1.			
DR	SMART: SM00406; IGV: 1.			
KM	Immunoglobulin V region.			
FT	NON_TER			
SQ	SEQUENCE 97 AA; 10661 MW; C23CB33FE5DA893 CRC64;			
Query Match 11.4%; Score 14; DB 1; Length 97;				
Best local Similarity 100.0%; Pred. NO. 1.8e-07;				
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	16 GSILKSCASGFTF 29			
DB	16 GSILKSCASGFTF 29			
RESULT 2				
ID	HV57_MOUSE	STANDARD;	PRT;	98 AA.
AC	P18528;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig heavy chain V region 6.96.			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/CJ;			
RX	MEDLINE=89279149; PubMed=2499654;			

RA Levy N.S., Malliero U.V., Lebeque S.G., Gearhart P.J.;  
 RT "Early onset of somatic mutation in immunoglobulin VH genes during  
 the primary immune response.";  
 RL J. Exp. Med. 169:2007-2019(1989).  
 CC -1- MISCLEANEOS: THIS SEQUENCE BELONGS TO THE VH183 SUBFAMILY.  
 DR PIR: JTO501: HWS96.  
 DR HSSP: P01772: 2PB4.  
 DR InterPro: IPR003006: Ig\_MHC.  
 DR InterPro: IPR003596: Ig\_V.  
 DR Pfam: PF00047: Ig\_1.  
 DR SMART: SM00406: IGV\_1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 98  
 SQ SEQUENCE 98 AA: 11007 MW: B8644FE92F8F95B CRC64:  
 Query Match 11.4%; Score 14; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GSLKISCAASGFTF 29  
 Db 16 GSLKISCAASGFTF 29  
 |||||

RESULT 3  
 ID HV3B\_HUMAN STANDARD; PRT; 114 AA.  
 AC P01763;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-II region WEA.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=83273707; PubMed=6410398;  
 RA Goni F., Frangione B.;  
 RT "Amino acid sequence of the Fv region of a human monoclonal IgM  
 (protein WEA) with antibody activity against 3,4-pyruvylated  
 galactose in Klebsiella polysaccharides K30 and K33.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).  
 CC -1- MISCLEANEOS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY  
 AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH  
 WALDENSTROM'S MACROGLOBULINEMIA.  
 DR PIR: A02046: M3HUM.  
 DR HSSP: P01772: 2PB4.  
 DR InterPro: IPR003006: Ig\_MHC.  
 DR InterPro: IPR003596: Ig\_V.  
 DR Pfam: PF00047: Ig\_1.  
 DR SMART: SM00406: IGV\_1.  
 KW Immunoglobulin V region.  
 FT MOD\_RES 1  
 FT NON\_TER 114  
 SQ SEQUENCE 114 AA: 12256 MW: D88294F8418A07B7 CRC64:  
 Query Match 11.4%; Score 14; DB 1; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 2e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 MNWVROAPGKLEW 47  
 Db 34 MNWVROAPGKLEW 47  
 |||||

RESULT 4  
 ID HV3T\_HUMAN STANDARD; PRT; 116 AA.  
 AC P01781;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ig heavy chain V-II region GAL.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=75059123; PubMed=4803843;  
 RA Watanabe S., Barnikol H.U., Horn J., Berttram J., Hilschmann N.;  
 RT "The primary structure of a monoclonal IgM-immunoglobulin  
 (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-  
 type), subgroup H III. Architecture of the complete IgM-molecule.";  
 RT Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).  
 RL [2]  
 RP REVISION TO 28-33.  
 RA Hilschmann N.;  
 CC Submitted (JUN-1975) to the PIR data bank.  
 CC -1- MISCLEANEOS: THIS KU CHAIN WAS ISOLATED FROM A WALDENSTROM'S  
 MACROGLOBULIN.  
 DR PIR: A02064: M3HUGL.  
 DR HSSP: P01772: 2PB4.  
 DR InterPro: IPR003006: Ig\_MHC.  
 DR InterPro: IPR003596: Ig\_V.  
 DR Pfam: PF00047: Ig\_1.  
 DR SMART: SM00406: IGV\_1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 116  
 SQ SEQUENCE 116 AA: 12730 MW: 2C67CA9AAMAA1282 CRC64:  
 Query Match 11.4%; Score 14; DB 1; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 2e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WVRQAPGKLEWVA 49  
 Db 36 WVRQAPGKLEWVA 49  
 |||||

RESULT 5  
 ID HV3O\_HUMAN STANDARD; PRT; 117 AA.  
 AC P01776;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-II region WAS.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=74142702; PubMed=4522793;  
 RA Capra J.D., Kenne J.M.;  
 RT "Variable region sequences of five human immunoglobulin heavy chains  
 of the VH3 subgroup: definitive identification of four heavy chain  
 hypervariable regions.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).  
 CC -1- MISCLEANEOS: THIS CHAIN WAS ISOLATED FROM AN UG1 MYELOMA  
 PROTEIN.  
 DR PIR: A02059: G1HUMS.  
 DR HSSP: P01772: 2PB4.  
 DR InterPro: IPR003006: Ig\_MHC.  
 DR InterPro: IPR003596: Ig\_V.  
 DR Pfam: PF00047: Ig\_1.  
 DR SMART: SM00406: IGV\_1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 117  
 SQ SEQUENCE 117 AA: 13091 MW: 201DEF0E1E53D9BF CRC64:  
 Query Match 11.4%; Score 14; DB 1; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;



Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WYRQAPGKLEMYA 49  
 DB 36 WYRQAPGKLEMYA 49

## RESULT 6

HV53\_MOUSE  
 ID HV53\_MOUSE STANDARD; PRT; 117 AA.  
 AC P18524;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V region RF precursor.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BAIB/CJ;  
 RX MEDLINE=89279149; PubMed=2499654;  
 RA Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;  
 RT "Early onset of somatic mutation in immunoglobulin VH genes during  
 the primary immune response.";  
 RL J. Exp. Med. 169:2007-2019(1989).  
 CC -I- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.  
 DR HSSP; J0503; HVMSRF.  
 DR HSSP; P01810; 2FBJ.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR InterPro: IPR003596; IG\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IgV; 1.  
 KW Immunoglobulin V region; Hybridoma; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 117 IG HEAVY CHAIN V REGION RF.  
 FT DOMAIN 20 49 FRAMEWORK-1.  
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
 FT DOMAIN 55 68 FRAMEWORK-2.  
 FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
 FT DOMAIN 86 117 FRAMEWORK-3.  
 FT DISULFID 41 115 BY SIMILARITY.  
 FT NON\_TER 117 117  
 SO SEQUENCE 117 AA; 12866 MW; 2CE3295F390F725B CRC64;

Query Match 11.4%; Score 14; DB 1; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GSKLSCAASGTF 29  
 DB 35 GSKLSCAASGTF 48

## RESULT 7

HV54\_MOUSE  
 ID HV54\_MOUSE STANDARD; PRT; 117 AA.  
 AC P18525;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V region 5-84 precursor.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BAIB/CJ;  
 RX MEDLINE=89279149; PubMed=2499654;  
 RA Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;  
 RT "Early onset of somatic mutation in immunoglobulin VH genes during

RT the primary immune response.";  
 RL J. Exp. Med. 169:2007-2019(1989).  
 CC -I- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.  
 DR PIR; J0505; HVMS84.  
 DR HSSP; P01810; 2FBJ.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR InterPro: IPR003596; IG\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IgV; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 117 IG HEAVY CHAIN V REGION 5-84.  
 FT DOMAIN 20 49 FRAMEWORK-1.  
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
 FT DOMAIN 55 68 FRAMEWORK-2.  
 FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
 FT DOMAIN 86 117 FRAMEWORK-3.  
 FT DISULFID 41 115 BY SIMILARITY.  
 FT NON\_TER 117 117  
 SO SEQUENCE 117 AA; 12872 MW; 234055CB6A469861 CRC64;

Query Match 11.4%; Score 14; DB 1; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GSKLSCAASGTF 29  
 DB 35 GSKLSCAASGTF 48

## RESULT 8

HV31\_HUMAN  
 ID HV31\_HUMAN STANDARD; PRT; 119 AA.  
 AC P01770;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region NIE.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC MEDLINE=77070269; PubMed=826475;  
 RA Ponslingl H., Hilschmann N.;  
 RT "The rule of antibody structure. The primary structure of a  
 monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The  
 chymotryptic peptides of the H-chain, alignment of the tryptic  
 peptides and discussion of the complete structure.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).  
 RN [2]  
 RP DISULFIDE BOND.  
 RX MEDLINE=77070267; PubMed=1002129;  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RT "Rule of antibody structure. The primary structure of a monoclonal  
 IgG1 immunoglobulin (myeloma protein NIE). I: Purification and  
 characterization of the protein, the L- and H-chains, the  
 cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976)  
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IG01 MYELOMA  
 PROTEIN.  
 CC PIR; A02053; GIHUNI.  
 DR HSSP; P01772; 2FBJ.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR InterPro: IPR003596; IG\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IgV; 1.  
 KW Immunoglobulin V region.  
 FT MOD\_RES 1 1  
 FT DISULFID 22 96 PYRROLIDONE CARBOXYLIC ACID.  
 FT NON\_TER 119 119  
 SO SEQUENCE 119 AA; 13242 MW; C96935A655E165B CH:54;

Query Match 11.4%; Score 14; DB 1; Length 119;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WVRQAPGKGLEWYA 49  
DB 36 WVRQAPGKGLEWYA 49

## RESULT 9

HV3M\_HUMAN STANDARD; PRT; 119 AA.  
ID HV3M\_HUMAN  
AC P01774;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 1g heavy chain V-III region F0M.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=75046755; PubMed=4139708;  
RA Capra J.D., Kehoe J.M.;  
RT "Structure of antibodies with shared idiotypy: the complete sequence  
RT of the heavy chain variable regions of two immunoglobulin M  
RT anti-gamma globulins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA  
CC GLOBULIN ACTIVITY.  
DR PIR: A02057; M3HUPM.  
DR HSSP: P01772; 2FBA.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig\_1.  
DR SMART: SM00406; IgV\_1.  
KW Immunoglobulin V region.  
FT VARIANT 54  
FT N -> D (PROBABLY DUE TO DEAMINATION  
FT DURING ISOLATION).  
FT /FTID=VAR\_003966.  
SQ SEQUENCE 119 AA; 12953 MW; 2E018AF4DCB2610 CRC64;

Query Match 11.4%; Score 14; DB 1; Length 119;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WVRQAPGKGLEWYA 49  
DB 36 WVRQAPGKGLEWYA 49

## RESULT 10

HV3N\_HUMAN STANDARD; PRT; 119 AA.  
ID HV3N\_HUMAN  
AC P01775;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 1g heavy chain V-III region LAY.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=75046755; PubMed=4139708;  
RA Capra J.D., Kehoe J.M.;  
RT "Structure of antibodies with shared idiotypy: the complete sequence  
RT of the heavy chain variable regions of two immunoglobulin M  
RT anti-gamma globulins.";

Query Match 11.4%; Score 14; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RL Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA  
CC GLOBULIN ACTIVITY.  
DR PIR: A02058; M3HULY.  
DR HSSP: P01772; 2FBA.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig\_1.  
DR SMART: SM00406; IgV\_1.  
KW Immunoglobulin V region.  
FT NON\_TER 119  
SQ SEQUENCE 119 AA; 12858 MW; D6338098794DCF5E CRC64;

Query Match 11.4%; Score 14; DB 1; Length 119;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WVRQAPGKGLEWYA 49  
DB 36 WVRQAPGKGLEWYA 49

## RESULT 11

HV3J\_HUMAN STANDARD; PRT; 121 AA.  
ID HV3J\_HUMAN  
AC P01771;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 1g heavy chain V-III region HIL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=79124695; PubMed=420800;  
RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;  
RT "Amino acid sequence of the VH region of human myeloma  
RT cryoimmunoglobulin IgG H11.";  
RL Biochemistry 18:553-560(1979).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA  
CC PROTEIN.  
DR PIR: A02054; GIHUL.  
DR HSSP: P01772; 2FBA.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig\_1.  
DR SMART: SM00406; IgV\_1.  
KW Immunoglobulin V region.  
FT MOD\_RES 1  
FT NON\_TER 121  
SQ SEQUENCE 121 AA; 13566 MW; 480FC53610EF5DAB CRC64;

Query Match 11.4%; Score 14; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WVRQAPGKGLEWYA 49  
DB 36 WVRQAPGKGLEWYA 49

## RESULT 12

HV3K\_HUMAN STANDARD; PRT; 126 AA.  
ID HV3K\_HUMAN  
AC P01772;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 1g heavy chain V-III region KOL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE, AND DISULFIDE BONDS.  
 RA MEDLINE=83289131; PubMed=6884994;  
 RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;  
 RT "Three-dimensional structure determination of antibodies. Primary  
 structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RA MEDLINE=81072295; PubMed=7441755;  
 RA Marguaret M., Delsenhofer J., Huber R., Palm W.;  
 RT "Crystallographic refinement and atomic models of the intact  
 immunoglobulin molecule and its antigen-binding fragment at 3.0 A  
 and 1.0-A resolution.";  
 RL J. Mol. Biol. 141:369-391(1980).  
 DR PIR: A02055; G1HUKL.  
 DR PDB: 2FB4; 12-JUL-89.  
 DR PDB: 2IG2; 12-JUL-89.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 KW Immunoglobulin V region; 3D-structure.  
 FT MOD\_RES 1 1  
 FT DISULFID 22 96 PYRROLIDONE CARBOXYLIC ACID.  
 FT DISULFID 105 110  
 FT STRAND 3 7  
 FT STRAND 11 12  
 FT TURN 14 15  
 FT STRAND 18 25  
 FT STRAND 29 31  
 FT HELIX 34 39  
 FT TURN 41 42  
 FT STRAND 46 51  
 FT TURN 53 54  
 FT STRAND 58 60  
 FT TURN 62 67  
 FT STRAND 68 73  
 FT TURN 74 77  
 FT STRAND 78 83  
 FT HELIX 88 90  
 FT STRAND 92 99  
 FT TURN 106 106  
 FT TURN 107 108  
 FT STRAND 109 109  
 FT STRAND 113 113  
 FT STRAND 120 124  
 FT NON\_TER 126 126  
 SQ SEQUENCE 126 AA; 13718 MW; EDD71B52B16F8776 CRC64;  
 Query Match 11.4%; Score 14; DB 1; Length 126;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 36 WYRQAPGKGLEWVA 49  
 DB 36 WYRQAPGKGLEWVA 49  
 RESULT 13  
 ID HV01\_RAT STANDARD; PRT; 142 AA.  
 AC P01805;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V region IR2 precursor.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=83064537; PubMed=6292865;  
 RA Hellman L., Petersson U., Engstrom A., Karlsson L., Bennich H.;  
 RT "Structure and evolution of the heavy chain from rat immunoglobulin  
 E.";  
 RL Nucleic Acids Res. 10:6041-6049(1982).  
 CC -1- MISCELLANEOUS: THE MRNA WAS ISOLATED FROM AN IGE-SECRETING  
 CC IMMUNOCYTOMA THAT ARISES SPONTANEOUSLY IN LOU/C/MSL RATS.  
 DR PIR: A02075; EVRTR2.  
 DR HSSP: P01789; 1MCP.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 142 IG HEAVY CHAIN V REGION IR2.  
 FT NON\_TER 142 142  
 SQ SEQUENCE 142 AA; 16024 MW; DE29BECFE74DF3B CRC64;  
 Query Match 11.4%; Score 14; DB 1; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 36 WYRQAPGKGLEWVA 49  
 DB 55 WYRQAPGKGLEWVA 68  
 RESULT 14  
 ID HV35\_HUMAN STANDARD; PRT; 115 AA.  
 AC P01780;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region JON.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=74142702; PubMed=4522793;  
 RA Capra J.D., Kehoe J.M.;  
 RT "Variable region sequences of five human immunoglobulin heavy chains  
 of the VH3 subgroup: definitive identification of four heavy chain  
 hypervariable regions.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).  
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG3 MYELOMA  
 CC PROTEIN.  
 DR PIR: A02063; G3HUNJ.  
 DR HSSP: P01772; 2FBA.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 115 115  
 SQ SEQUENCE 115 AA; 12563 MW; 68B668B531C12514 CRC64;  
 Query Match 10.6%; Score 13; DB 1; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 36 WYRQAPGKGLEWVA 48  
 DB 36 WYRQAPGKGLEWVA 48  
 RESULT 15  
 ID HV30\_HUMAN

```

ID HV30_HUMAN STANDARD; PRT; 116 AA.
AC P01778;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region ZAP.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kehoe J.W.;
RT "Variable region sequences of five human immunoglobulin heavy chains
RT of the VH3 subgroup: definitive identification of four heavy chain
RT hypervariable regions."
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGAI MYELOMA
CC PROTEIN.
DR PIR: A02061; A1H02P.
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12582 MW; 892F8C217CEC9865 CRC64;

Query Match 10.6%; Score 13; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. NO. 2.3e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 36 WVRQAPGKGLWV 48
DB 36 WVRQAPGKGLWV 48

```

Search completed: November 27, 2002, 07:28:47  
 Job time : 12.8726 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:24:25 ; Search time 34.668 Seconds  
(without alignments)  
731.044 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 123  
Sequence: 1 EVMVSGGGLVQPKGSLKL.....SGIDYAMQYNGQTSITVSS 123

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 671580 segs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.todent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	13.8	494	4 Q96K68	Q96K68 homo sapien
2	17	13.8	597	4 Q96B89	Q96B89 homo sapien
3	15	12.2	118	4 Q9UL91	Q9UL91 homo sapien
4	15	12.2	471	4 Q8TC77	Q8TC77 homo sapien
5	14	11.4	95	4 Q9UL86	Q9UL86 homo sapien
6	14	11.4	113	4 Q9UL90	Q9UL90 homo sapien
7	14	11.4	116	4 Q9UL93	Q9UL93 homo sapien
8	14	11.4	119	11 Q920E7	Q920E7 mus musculu
9	14	11.4	122	4 Q9UL84	Q9UL84 homo sapien
10	14	11.4	437	11 Q9RIA4	Q9RIA4 mus musculu
11	14	11.4	487	11 Q99KA4	Q99KA4 mus musculu
12	14	11.4	613	4 Q8WU1	Q8WU1 homo sapien
13	13	10.6	104	4 Q9UL87	Q9UL87 homo sapien
14	13	10.6	112	4 Q9HCL1	Q9HCL1 homo sapien
15	13	10.6	112	4 Q9UGP3	Q9UGP3 homo sapien
16	13	10.6	118	4 Q9UL72	Q9UL72 homo sapien

17	13	10.6	121	4 Q9UL71	Q9UL71 homo sapien
18	13	10.6	131	4 Q9UL88	Q9UL88 homo sapien
19	13	10.6	573	4 Q8WU38	Q8WU38 homo sapien
20	12	9.8	124	6 Q9N0W6	Q9N0W6 oryctolagus
21	12	9.8	124	6 Q9N0W4	Q9N0W4 oryctolagus
22	12	9.8	143	11 Q924R0	Q924R0 mus musculu
23	12	9.8	145	11 Q924R4	Q924R4 mus musculu
24	12	9.8	146	11 Q924Q3	Q924Q3 mus musculu
25	12	9.8	298	11 Q9QYF0	Q9QYF0 mus musculu
26	11	8.9	15	11 Q9QV16	Q9QV16 ratu sp.
27	11	8.9	110	11 Q9JL77	Q9JL77 mus musculu
28	11	8.9	117	11 Q9QXE9	Q9QXE9 mus musculu
29	11	8.9	121	11 Q99NG4	Q99NG4 mus musculu
30	11	8.9	145	11 Q924R1	Q924R1 mus musculu
31	11	8.9	145	11 Q924Q9	Q924Q9 mus musculu
32	11	8.9	145	11 Q924Q6	Q924Q6 mus musculu
33	11	8.9	145	11 Q924P7	Q924P7 mus musculu
34	11	8.9	426	11 Q9DCD9	Q9DCD9 mus musculu
35	11	8.9	469	11 Q8R3V9	Q8R3V9 mus musculu
36	11	8.9	473	11 Q91Z05	Q91Z05 mus musculu
37	11	8.9	484	11 Q91LA6	Q91LA6 mus musculu
38	11	8.9	484	11 Q8VEA0	Q8VEA0 mus musculu
39	11	8.1	142	11 Q924Q1	Q924Q1 mus musculu
40	10	8.1	147	4 Q9Y509	Q9Y509 homo sapien
41	10	8.1	150	4 Q9Y298	Q9Y298 homo sapien
42	10	8.1	482	11 Q91X82	Q91X82 mus musculu
43	10	8.1	486	11 Q91Z07	Q91Z07 mus musculu
44	9	7.3	64	11 Q61750	Q61750 mus musculu
45	9	7.3	119	5 Q9GY22	Q9GY22 schistosoma

## ALIGNMENTS

RESULT 1  
ID Q96K68 PRELIMINARY: PRT: 494 AA.  
AC Q96K68: 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE CDNA FLJ14473 fis, clone MAMMA1001080, highly similar to Homo sapiens SNC73 protein (SNC73) mRNA.  
DE Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP TISSUE=AMMARY GLAND;  
RC Itogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Ninomiya K., Iwayanagi T.;  
RA "NEO human cDNA sequencing project."  
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AK027379; BAB55072.1; -  
DR InterPro: IPR003006; Ig\_MHC.  
DR Pfam: PF00047; Ig\_4.  
DR PROSITE: PS00290; Ig\_MHC; UNKNOWN.1.  
SQ SEQUENCE 494 AA; 5308 MW; 9A1D7AE5AEAC0E CkC64;

Query Match 13.8%; Score 17; DB 4; Length 494;

Best local similarity 100.0%; Pred. No. 4.6e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 YAMNWVROAPGKGLVW 48  
Db 51 YAMNWVROAPGKGLVW 67

## RESULT 2

Q96BB9 PRELIMINARY; PRT; 597 AA.  
 ID 096BB9  
 AC 096BB9;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical 65.0 kDa protein.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=B-CELL;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC015760; AHI5760.1;  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig\_5.  
 DR PROSITE: PS00290; IG\_MHC; UNKNOWN\_3.  
 KW Hypothetical protein.  
 SQ SEQUENCE 597 AA; 65039 MW; 4FCA3AD8EC263D9 CRC64;

Query Match 13.8%; Score 17; DB 4; Length 597;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 YAMNVRQAPGKLEWV 48  
 DB 51 YAMNVRQAPGKLEWV 67

## RESULT 3

Q9UL91 PRELIMINARY; PRT; 118 AA.  
 ID 09UL91  
 AC 09UL91;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Myosin-reactive immunoglobulin heavy chain variable region  
 DE (Fragment).  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RL MEDLINE=98277139; PubMed=9614934;  
 RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Beiney S.M.,  
 RA Young D.C.;  
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
 RT fetus";  
 RT Clin. Immunol. Immunopathol. 87:184-192(1998).  
 DR EMBL: AF035023; AAD56259.1;  
 DR HSSP: P01772; 2PB4.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_1.  
 DR SMART: SM00406; IGV\_1.  
 FT NON\_TER 1  
 FT NON\_TER 118  
 SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;

Query Match 12.2%; Score 15; DB 4; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 MNWVRQAPGKLEWV 48  
 DB 34 MNWVRQAPGKLEWV 48

## RESULT 4

Q8TC77 PRELIMINARY; PRT; 471 AA.  
 ID 08TC77  
 AC 08TC77;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical 51.8 kDa protein.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SPLEEN;  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC024289; AAH24289.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;

Query Match 12.2%; Score 15; DB 4; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 MNWVRQAPGKLEWV 48  
 DB 53 MNWVRQAPGKLEWV 67

## RESULT 5

Q9ULB6 PRELIMINARY; PRT; 95 AA.  
 ID 09ULB6  
 AC 09ULB6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Immunoglobulin heavy chain (Fragment).  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tange Y., Kayano H.;  
 RT "Human VH gene sequence";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB035268; BAA87067.1;  
 DR HSSP: P01772; 2PB4.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_1.  
 DR SMART: SM00406; IGV\_1.  
 FT NON\_TER 1  
 FT NON\_TER 95  
 SQ SEQUENCE 95 AA; 10527 MW; 90A8C6D1D22574A CRC64;

Query Match 11.4%; Score 14; DB 4; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 WVRQAPGKLEWVA 49  
 DB 35 WVRQAPGKLEWVA 48

## RESULT 6

Q9UL90 PRELIMINARY; PRT; 113 AA.  
 ID 09UL90  
 AC 09UL90;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98277139; PubMed-9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
RL EMBL: AF035024; AAD56260.1; -.
DR HSSP: P01772; 2F84.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igv; 1.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match
Best Local Similarity 11.4%; Score 14; DB 4; Length 113;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLGVYA 49
DB 36 WVRQAPGKGLGVYA 49

RESULT 7
Q9UL93 PRELIMINARY; PRT; 116 AA.
ID Q9UL93;
AC Q9UL93;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98277139; PubMed-9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
RL EMBL: AF035021; AAD56257.1; -.
DR HSSP: P01772; 2F84.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igv; 1.
FT NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 12434 MW; ODA0348154DD6061 CRC64;

Query Match
Best Local Similarity 11.4%; Score 14; DB 4; Length 116;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLGVYA 49
DB 35 WVRQAPGKGLGVYA 48

```

```

RESULT 8
Q920E7 PRELIMINARY; PRT; 119 AA.
ID Q920E7;
AC Q920E7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Pterin-mimicking anti-Idiotope heavy chain variable region
DE (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Atkin J.D., Iape A., Jennings I.G., Horvath O., Cotton R.G.H.;
RT "Definition of the idiotope of Pterin-Mimicking Antibodies Expressed
RT in Mammalian Cells."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF307937; AAL09421.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
FT NON_TER 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 13025 MW; F6E90404381CA7C CRC64;

Query Match
Best Local Similarity 11.4%; Score 14; DB 11; Length 119;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GSKLSCAASGTF 29
DB 16 GSKLSCAASGTF 29

RESULT 9
Q9UL84 PRELIMINARY; PRT; 122 AA.
ID Q9UL84;
AC Q9UL84;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98277139; PubMed-9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
RL EMBL: AF035030; AAD56266.1; -.
DR HSSP: P01772; 2F84.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igv; 1.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 13579 MW; 360540A1366545B8 CRC64;

Query Match
Best Local Similarity 11.4%; Score 14; DB 4; Length 122;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLGVYA 49

```

Db 36 WVRQAPGKGLEWVA 49

## RESULT 10

Q9RLA4 PRELIMINARY; PRT; 437 AA.  
 AC Q9RLA4;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Gamma heavy chain of Mab7 (Fragment).  
 GN IGH-4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;  
 RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal  
 antibody (Mab 7, its light and heavy chains) and construction of a  
 single chain antibody (scFv).";  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF152372; AAD40243.1;  
 DR HSSP; P01842; 7FAB.  
 DR MCD; MGI:96446; Igh-4.  
 DR InterPro: IPR003600; Ig\_1like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 4.  
 DR SMART; SM00406; Igv; 1.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
 FT NON\_TER 1  
 FT SEQUENCE 437 AA; 48142 MW; 5C3A7BB3E7D697C CRC64;

Query Match 11.4%; Score 14; DB 11; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GSKLSCAASGTF 29

Db 15 GSKLSCAASGTF 28

## RESULT 11

Q99KA4 PRELIMINARY; PRT; 487 AA.  
 AC Q99KA4;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical 52.6 kDa protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC004786; AAH04786.1;  
 DR HSSP; P01810; 2FBJ.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003597; Ig\_C1.  
 DR InterPro: IPR003600; Ig\_1like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 4.  
 DR SMART; SM00409; Igv; 3.  
 DR SMART; SM00407; Igc1; 3.

DR SMART; SM00406; Igv; 1.  
 DR SMART; SM00410; Ig\_1like; 1.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 487 AA; 52554 MW; 7DC8E96DB333077B CRC64;

Query Match 11.4%; Score 14; DB 11; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GSKLSCAASGTF 29

Db 35 GSKLSCAASGTF 48

## RESULT 12

Q8WUK1 PRELIMINARY; PRT; 613 AA.  
 AC Q8WUK1;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical 67.3 kDa protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TONSIL;  
 RA Strausberg R.;  
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC020240; AAH20240.1;  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003597; Ig\_C1.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 5.  
 DR SMART; SM00409; Igv; 2.  
 DR SMART; SM00407; Igc1; 4.  
 DR SMART; SM00406; Igv; 1.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_3.  
 KW Hypothetical protein.  
 SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;

Query Match 11.4%; Score 14; DB 4; Length 613;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLEWVA 49

Db 55 WVRQAPGKGLEWVA 68

## RESULT 13

Q9UL87 PRELIMINARY; PRT; 104 AA.  
 AC Q9UL87;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Myosin-reactive immunoglobulin heavy chain variable region  
 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9827139; PubMed=9614934;  
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Betney S.M.,  
 Young D.C.;  
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal



```

RT fetus.;
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035027; AAD56263.1; -.
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Ig; 1.
FT NON_TER 1
FT NON_TER 104
SQ SEQUENCE 104 AA; 11598 MW; 611D1A3F40E96E7B CRC64;

Query Match
Best Local Similarity 100.0%; Score 13; DB 4; Length 104;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLEWV 48
DB 15 WVRQAPGKGLEWV 27

RESULT 14
Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Single chain Fv (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RT "An antibody fragment specific for native lysozyme: Isolation from a
RT human synthetic phage display library and characterization.";
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB049915; BAB16829.1; -.
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00406; Ig; 1.
DR SMART: SM00410; Ig_Like; 1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match
Best Local Similarity 100.0%; Score 13; DB 4; Length 112;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLEWV 48
DB 36 WVRQAPGKGLEWV 48

RESULT 15
Q9UGP3 PRELIMINARY; PRT; 112 AA.
AC Q9UGP3;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Immunoglobulin heavy chain (Fragment).
GN IGH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Krambovitis E.;
RA "Induction of somatic mutations in human B cells by in vitro
RT immunization.";
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ132560; CAB65078.1; -.
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Ig; 1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12226 MW; 0A1F17082C7A8CE3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 13; DB 4; Length 112;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLEWV 48
DB 19 WVRQAPGKGLEWV 31

Search completed: November 27, 2002, 07:30:26
Job time : 36.918 secs

```

1K (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:26:06 : Search time 17.0965 Seconds  
(without alignments)  
211.682 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 123

Sequence: 1 EVMILVSGSGLVQPKGSLKL.....SGIDYAMDYGGGTSLTVSS 123

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size: 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCrus.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	32.5	110	3	US-08-767-128-24
2	40	32.5	115	3	US-08-767-128-36
3	28	22.8	119	1	US-08-442-542-2
4	28	22.8	119	3	US-08-765-469-2
5	28	22.8	599	1	US-08-442-542-18
6	28	22.8	599	3	US-08-765-469-18
7	17	13.8	20	1	US-08-050-113-1
8	17	13.8	113	5	US-08-428-197-46
9	17	13.8	113	5	PCT-US93-10555-46
10	16	13.0	120	2	US-07-934-373C-2
11	16	13.0	120	2	US-07-934-373C-42
12	16	13.0	120	2	US-07-934-373C-44
13	16	13.0	120	2	US-07-934-373C-48
14	16	13.0	120	3	US-08-437-642B-2
15	16	13.0	120	3	US-08-437-642B-42
16	16	13.0	120	3	US-08-437-642B-44
17	16	13.0	120	4	US-08-146-206C-2
18	16	13.0	120	5	PCT-US93-07832-2
19	16	13.0	122	2	US-07-934-373C-20
20	16	13.0	122	2	US-07-934-373C-45
21	16	13.0	122	3	US-08-437-642B-20
22	16	13.0	122	3	US-08-437-642B-45
23	16	13.0	122	4	US-08-146-206C-20
24	16	13.0	122	4	US-08-146-206C-26
25	16	13.0	122	5	PCT-US93-07832-20
26	16	13.0	449	4	US-09-679-397-2
27	16	13.0	449	4	US-09-680-148-2

28	16	13.0	461	2	US-08-463-587A-26	Sequence 26, Appl
29	16	13.0	461	2	US-08-463-667A-4	Sequence 4, Appl
30	16	13.0	461	3	US-08-923-85A-26	Sequence 26, Appl
31	16	13.0	461	5	PCT-US91-09133-27	Sequence 27, Appl
32	15	12.2	98	1	US-08-211-202-116	Sequence 116, App
33	15	12.2	98	2	US-08-665-202-31	Sequence 31, Appl
34	15	12.2	109	2	US-08-379-057-32	Sequence 32, Appl
35	15	12.2	111	1	US-08-211-202-134	Sequence 134, Appl
36	15	12.2	112	1	US-08-211-202-133	Sequence 133, App
37	15	12.2	116	2	US-08-428-197-10	Sequence 10, Appl
38	15	12.2	116	2	US-09-027-449-50	Sequence 50, Appl
39	15	12.2	116	3	US-09-184-658-48	Sequence 48, Appl
40	15	12.2	116	3	US-08-804-444A-50	Sequence 50, Appl
41	15	12.2	116	4	US-09-026-985-50	Sequence 50, Appl
42	15	12.2	116	4	US-09-121-952A-50	Sequence 50, Appl
43	15	12.2	116	4	US-09-234-340A-50	Sequence 50, Appl
44	15	12.2	116	5	PCT-US93-10555-10	Sequence 10, Appl
45	15	12.2	117	3	US-08-545-809A-107	Sequence 107, App

#### ALIGNMENTS

RESULT 1  
US-08-767-128-24  
Sequence 24, Application US/08767128  
Patent No. 6111079  
GENERAL INFORMATION:  
APPLICANT: WYLLIE, DWANE E.  
APPLICANT: LOPEZ, OSVALDO  
APPLICANT: MURRAY, PETER JOSEPH  
APPLICANT: GOEBEL, PETER  
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 6111079west Center, 90 South Seventh St  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,128  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE: 04-DEC-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09258  
FILING DATE: 05-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/541,373  
FILING DATE: 10-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/462,798  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Carter, Charles G.  
REGISTRATION NUMBER: 35,093  
REFERENCE/DOCKET NUMBER: 8648,49USF1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612/371-5278  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
US-08-767-128-24

Query Match  
Best Local Similarity 100.0%; Pred. No. 1.4e-30;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 YADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVR 101  
DB 57 YADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVR 96

RESULT 2  
US-08-767-128-36  
Sequence 36, Application US/08767128  
Patent No. 6111079  
GENERAL INFORMATION:  
APPLICANT: WYLIE, DWANE E.  
APPLICANT: LOPEZ, OSVALDO  
APPLICANT: MURRAY, PETER JOSEPH  
APPLICANT: GOEBEL, PETER  
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Meltzer & Schmidt  
STREET: 3100 No. 6111079west Center, 90 South seventh St  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,128  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE: 04-DEC-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09258  
FILING DATE: 05-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/541,373  
FILING DATE: 10-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/462,798  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Carter, Charles G.  
REGISTRATION NUMBER: 35,093  
REFERENCE/DOCKET NUMBER: 8648.49USF1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612/371-5278  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ. ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
US-08-767-128-36

Query Match  
Best Local Similarity 100.0%; Pred. No. 1.5e-30;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 YADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVR 101  
DB 62 YADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVR 101

RESULT 3  
US-08-442-542-2  
Sequence 2, Application US/08442542  
Patent No. 5686600  
GENERAL INFORMATION:  
APPLICANT: Carozzi, Nadine B.  
APPLICANT: Koziel, Michael G.  
TITLE OF INVENTION: Antibodies which bind to Insect Gut  
TITLE OF INVENTION: Proteins and their use  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,542  
FILING DATE: 16-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/267,641  
FILING DATE: 28-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-442-542-2

Query Match  
Best Local Similarity 100.0%; Pred. No. 3e-19;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AMNVRQAPGKGLGEMVARIKSNMYAT 60  
DB 33 AMNVRQAPGKGLGEMVARIKSNMYAT 60

RESULT 4

US-08-765-469-2  
; Sequence 2, Application US/08765469  
; Patent No. 6069301  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine B.  
; APPLICANT: Koziel, Michael G.  
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
; TITLE OF INVENTION: Proteins and their Use  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ciba-Geigy Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/267,641  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sprull, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: CGC 1750  
; TELEPHONE: 919-541-8615  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-765-469-2

Query Match 22.8%; Score 28; DB 3; Length 119;  
Best Local Similarity 100.0%; Pred. No. 3e-19;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AMNWVQAPGKGLEWVARIRSKSNYYAT 60  
|||||  
DB 33 AMNWVQAPGKGLEWVARIRSKSNYYAT 60

RESULT 5  
US-08-442-542-18  
; Sequence 18, Application US/08442542  
; Patent No. 5686600  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine B.  
; APPLICANT: Koziel, Michael G.  
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
; TITLE OF INVENTION: Proteins and their Use  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ciba-Geigy Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/442,542  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/267,641  
; FILING DATE: 28-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sprull, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: CGC 1750  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8615  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 599 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-442-542-18

Query Match 22.8%; Score 28; DB 1; Length 599;  
Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AMNWVQAPGKGLEWVARIRSKSNYYAT 60  
|||||  
DB 187 AMNWVQAPGKGLEWVARIRSKSNYYAT 214

RESULT 6  
US-08-765-469-18  
; Sequence 18, Application US/08765469  
; Patent No. 6069301  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine B.  
; APPLICANT: Koziel, Michael G.  
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
; TITLE OF INVENTION: Proteins and their Use  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ciba-Geigy Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/765,469  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/267,641  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sprull, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: CGC 1750  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8615  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 599 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
US-08-765-469-18
Query Match 22.8%; Score 28; DB 3; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.2e-18;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AMNVROAPGKGLVWIRIRSKSNYYAT 60
|||||
Db 187 AMNVROAPGKGLVWIRIRSKSNYYAT 214

RESULT 7
US-08-050-113-1
; Sequence 1, Application US/08050113
; Patent No. 5514554
; GENERAL INFORMATION:
; APPLICANT: Bacus, Sarah S
; APPLICANT: Yarden, Yosef
; APPLICANT: Sela, Michael
; TITLE OF INVENTION: Methods and Compositions for Cancer
; TITLE OF INVENTION: Therapy and for Prognosticating Responses to Cancer
; TITLE OF INVENTION: Therapy
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton, Dickinson &
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07417-1880
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,113
; FILING DATE: 07-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hightet, David W
; REGISTRATION NUMBER: 30,265
; REFERENCE/DOCKET NUMBER: P-2624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 847 5317
; TELEFAX: 201 848 9228
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-050-113-1

Query Match 13.8%; Score 17; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LVESGGGLVQPKGSLKL 20
|||||
Db 4 LVESGGGLVQPKGSLKL 20

RESULT 8
US-08-428-197-46
; Sequence 46, Application US/08428197
; Patent No. 5891438
; GENERAL INFORMATION:
; APPLICANT: SILVERMAN, GREGG J.
; TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
```

```
; TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
; TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,197
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: PCT/US93/10555
; APPLICATION NUMBER: PCT/US93/10555
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD-2630
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-428-197-46

Query Match 13.8%; Score 17; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AMNVROAPGKGLVW 49
|||||
Db 33 AMNVROAPGKGLVW 49

RESULT 9
PCT-US93-10555-46
; Sequence 46, Application PC/TUS9310555
; GENERAL INFORMATION:
; APPLICANT: SILVERMAN, GREGG J.
; TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
; TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
; TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10555
; FILING DATE: 29-OCT-1993
; CLASSIFICATION:
```

ATTORNEY/AGENT INFORMATION:  
NAME: Howells, Stacy L.  
REGISTRATION NUMBER: 34,842  
REFERENCE/DOCKET NUMBER: FD-2630  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 113 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-10555-46

Query Match 13.8%; Score 17; DB 5; Length 113;  
Best Local Similarity 100.0%; Pred. No. 6.5e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 33 AMNVROAPGKGLEWVA 49  
|||||  
Db 33 AMNVROAPGKGLEWVA 49

## RESULT 10

US-07-934-373C-2  
Sequence 2, Application US/07934373C  
Patent No. 5821337

GENERAL INFORMATION:  
APPLICANT: Paul J. Carter  
APPLICANT: Leonard G. Presta  
TITLE OF INVENTION: Immunoglobulin Variants  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/934,373C  
FILING DATE: 21-Aug-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION NUMBER: PCT/US92/05126

FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/715272  
FILING DATE: 14-JUN-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0709P2  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-07-934-373C-2

Query Match 13.0%; Score 16; DB 2; Length 120;  
Best Local Similarity 100.0%; Pred. No. 6e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 36 WVRQAPGKGLEWVARI 51  
|||||  
Db 36 WVRQAPGKGLEWVARI 51

## RESULT 11

US-07-934-373C-42  
Sequence 42, Application US/07934373C  
Patent No. 5821337

GENERAL INFORMATION:  
APPLICANT: Paul J. Carter  
APPLICANT: Leonard G. Presta  
TITLE OF INVENTION: Immunoglobulin Variants  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/934,373C  
FILING DATE: 21-Aug-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION NUMBER: PCT/US92/05126

FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/715272  
FILING DATE: 14-JUN-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0709P2  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-07-934-373C-42

Query Match 13.0%; Score 16; DB 2; Length 120;  
Best Local Similarity 100.0%; Pred. No. 6e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 36 WVRQAPGKGLEWVARI 51  
|||||  
Db 36 WVRQAPGKGLEWVARI 51

## RESULT 12

US-07-934-373C-44  
Sequence 44, Application US/07934373C  
Patent No. 5821337

GENERAL INFORMATION:  
APPLICANT: Paul J. Carter  
APPLICANT: Leonard G. Presta  
TITLE OF INVENTION: Immunoglobulin Variants  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California





Best Local Similarity 100.0%; Pred. No. 6e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 36 WYRQAPGKGLEWVARI 51  
Db 36 WYRQAPGKGLEWVARI 51

RESULT 15

US-08-437-642B-42  
; Sequence 42, Application US/08437642B  
; Patent No. 6054297  
; GENERAL INFORMATION:  
; APPLICANT: Paul J. Carter  
; APPLICANT: Leonard G. Presta  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437.642B  
; FILING DATE: 09-May-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934373  
; FILING DATE: 21-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709p2C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-437-642B-42

Query Match 13.0%; Score 16; DB 3; Length 120;  
Best Local Similarity 100.0%; Pred. No. 6e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 WYRQAPGKGLEWVARI 51  
Db 36 WYRQAPGKGLEWVARI 51

Search completed: November 27, 2002, 07:32:10  
Job time : 18.0965 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:28:06 ; Search time 11.3977 Seconds  
(without alignments)  
171.849 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 123

Sequence: 1 EVMLVSGGGLVQPKGSLK.....SGIDYAMDYWGQTSLSLTVSS 123

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 102317 seqs, 15924203 residues

Word size : 0

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	100.0	123	10	US-09-893-615-87
2	38	30.9	100	10	US-09-840-459-35
3	33	26.8	143	10	US-09-881-823-16
4	24	19.5	101	10	US-09-840-459-34
5	24	19.5	117	10	US-09-835-087-8
6	24	19.5	117	10	US-09-809-739-12
7	24	19.5	117	10	US-09-840-459-10
8	24	19.5	148	10	US-09-840-459-100
9	21	17.1	117	10	US-09-835-087-12
10	21	17.1	117	10	US-09-835-087-13
11	21	17.1	117	10	US-09-809-739-22
12	21	17.1	117	10	US-09-809-739-23
13	21	17.1	117	10	US-09-840-459-19
14	21	17.1	117	10	US-09-840-459-20
15	17	13.8	116	10	US-09-828-708-12
16	17	13.8	117	10	US-09-835-087-10
17	17	13.8	117	10	US-09-835-087-11
18	17	13.8	117	10	US-09-809-739-20
19	17	13.8	117	10	US-09-809-739-21

20	17	13.8	117	10	US-09-840-459-17	Sequence 17, Appl
21	17	13.8	117	10	US-09-840-459-18	Sequence 18, Appl
22	17	13.8	119	10	US-09-840-459-104	Sequence 104, Appl
23	16	13.0	120	10	US-09-971-543-11	Sequence 11, Appl
24	16	13.0	252	10	US-09-971-543-1	Sequence 1, Appl
25	15	12.2	98	12	US-10-066-895-4	Sequence 4, Appl
26	15	12.2	103	10	US-09-864-761-47355	Sequence 47355, A
27	15	12.2	115	10	US-09-974-449-2	Sequence 2, Appl
28	15	12.2	118	10	US-09-056-1608-108	Sequence 108, Appl
29	15	12.2	118	10	US-09-056-1608-112	Sequence 112, Appl
30	15	12.2	118	10	US-09-056-1608-116	Sequence 116, Appl
31	15	12.2	118	10	US-09-056-1608-118	Sequence 118, Appl
32	15	12.2	123	10	US-09-056-1608-7	Sequence 7, Appl
33	15	12.2	123	10	US-09-056-1608-14	Sequence 14, Appl
34	15	12.2	123	10	US-09-056-1608-127	Sequence 127, Appl
35	15	12.2	128	10	US-09-850-165-81	Sequence 81, Appl
36	15	12.2	128	10	US-09-850-165-110	Sequence 110, Appl
37	15	12.2	218	10	US-09-974-449-35	Sequence 35, Appl
38	14	11.4	14	10	US-09-253-794-33	Sequence 33, Appl
39	14	11.4	14	10	US-09-828-708-64	Sequence 64, Appl
40	14	11.4	87	10	US-09-840-459-53	Sequence 53, Appl
41	14	11.4	88	10	US-09-840-459-51	Sequence 51, Appl
42	14	11.4	89	10	US-09-840-459-48	Sequence 48, Appl
43	14	11.4	89	10	US-09-840-459-49	Sequence 49, Appl
44	14	11.4	98	10	US-09-840-459-38	Sequence 38, Appl
45	14	11.4	98	10	US-09-840-459-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1

US-09-893-615-87

; Sequence 87, Application US/09893615

; Patent No. US20020082395A1

; GENERAL INFORMATION:

; APPLICANT: Fischer, Gerald W.

; Schuman, Richard F.

; Wong, Hing

; Stinson, Jeffrey L.

; TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND

; CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR

; POSITIVE BACTERIA

; NUMBER OF SEQUENCES: 89

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &

; CITY: Washington

; STATE: DC

; STREET: 1300 I Street, NW

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC Compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/893,615

; FILING DATE: 29-Jun-2001

; ATTORNEY/AGENT INFORMATION:

; NAME: Elnaudi, Carol P.

; REGISTRATION NUMBER: 32,220

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-408-4000

; TELEFAX: 202-408-4400

; INFORMATION FOR SEQ ID NO: 87:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 123 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 87:
US-09-893-615-87
Query Match      100.0%; Score 123; DB 10; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.5e-103; Mismatches 0; Indels 0; Gaps 0;
Matches 123; Conservative 0;

QY 1 EYMLVSGGLVQPKGSLKSLCAASGFTNNYAMNVRQAPGKGLWVARIRSKSNYYAT 60
DB 1 EYMLVSGGLVQPKGSLKSLCAASGFTNNYAMNVRQAPGKGLWVARIRSKSNYYAT 60

QY 61 FYADSVKDRFTISRDDSQSMLYLQMNMLKTEDTAMYCYVRRGASGIDYAMDYWGQTSLT 120
DB 61 FYADSVKDRFTISRDDSQSMLYLQMNMLKTEDTAMYCYVRRGASGIDYAMDYWGQTSLT 120

QY 121 VSS 123
DB 121 VSS 123

RESULT 2
US-09-840-459-35
; Sequence 35, Application US/09840459
; Patent No. US20020150576A1
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Slobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-840-459-35

Query Match      30.9%; Score 38; DB 10; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 YADSVKDRFTISRDDSQSMLYLQMNMLKTEDTAMYCYV 99
DB 62 YADSVKDRFTISRDDSQSMLYLQMNMLKTEDTAMYCYV 99

RESULT 3
US-09-881-823-16
; Sequence 16, Application US/09881823
; Patent No. US2002008066A1
; GENERAL INFORMATION:
; APPLICANT: SHI, WENYUAN
; APPLICANT: ANDERSON, MAXWELL
; APPLICANT: MORRISON, SHERIE
; APPLICANT: TRINH, RYAN
; APPLICANT: WIMS, LETITIA
; APPLICANT: CHEN, LI
; TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries

; FILE REFERENCE: 22851-032
; CURRENT APPLICATION NUMBER: US/09/881,823
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 07/378,577
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Murine
US-09-881-823-16

Query Match      26.8%; Score 33; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 9.2e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 DRFTISRDDSQSMLYLQMNMLKTEDTAMYCYV 100
DB 87 DRFTISRDDSQSMLYLQMNMLKTEDTAMYCYV 119

RESULT 4
US-09-840-459-34
; Sequence 34, Application US/09840459
; Patent No. US20020150576A1
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Slobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-840-459-34

Query Match      19.5%; Score 24; DB 10; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LVESGGGLVQPKGSLKSLCAASGF 27
DB 4 LVESGGGLVQPKGSLKSLCAASGF 27

RESULT 5
US-09-835-087-8
; Sequence 8, Application US/09835087
; Patent No. US20020042370A1
; GENERAL INFORMATION:
; APPLICANT: Wayne W. Hancock
; TITLE OF INVENTION: Method of Treating Graft Rejection Using
; TITLE OF INVENTION: Inhibitors of CCR2 Function
; FILE REFERENCE: 1855.2008-003
; CURRENT APPLICATION NUMBER: US/09/835,087
```

```
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/549,448
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)...(117)
; OTHER INFORMATION: Murine mAb 1D9 heavy chain variable region
US-09-835-087-8

Query Match          19.5%; Score 24; DB 10; Length 117;
Best Local Similarity 100.0%; Pred. No. 8.3e-15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LVESGGGLVQPKGSLKLSCAASGF 27
   |||||
Db 4 LVESGGGLVQPKGSLKLSCAASGF 27

RESULT 6
US-09-809-739-12
; Sequence 12, Application US/09809739
; Patent No. US20020106369A1
; GENERAL INFORMATION:
; APPLICANT: Horvath, Christopher J.
; APPLICANT: Rao, Patricia E.
; TITLE OF INVENTION: Method of Inhibiting Stenosis and
; FILE REFERENCE: 1855.1069-003
; CURRENT APPLICATION NUMBER: US/09/809,739
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: US 09/528,267
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)...(117)
; OTHER INFORMATION: Murine mAb 1D9 heavy chain variable region
; NAME/KEY: SITE
; LOCATION: (31)...(35)
; OTHER INFORMATION: CDR1
; NAME/KEY: SITE
; LOCATION: (50)...(68)
; OTHER INFORMATION: CDR2
; NAME/KEY: SITE
; LOCATION: (101)...(106)
; OTHER INFORMATION: CDR3
; OTHER INFORMATION: Mouse
US-09-809-739-12

Query Match          19.5%; Score 24; DB 10; Length 117;
Best Local Similarity 100.0%; Pred. No. 8.3e-15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LVESGGGLVQPKGSLKLSCAASGF 27
   |||||
Db 4 LVESGGGLVQPKGSLKLSCAASGF 27

RESULT 7
US-09-840-459-10
; Sequence 10, Application US/09840459
; Patent No. US20020150576A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-840-459-10

Query Match          19.5%; Score 24; DB 10; Length 117;
Best Local Similarity 100.0%; Pred. No. 8.3e-15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LVESGGGLVQPKGSLKLSCAASGF 27
   |||||
Db 4 LVESGGGLVQPKGSLKLSCAASGF 27

RESULT 8
US-09-840-459-100
; Sequence 100, Application US/09840459
; Patent No. US20020150576A1
; GENERAL INFORMATION:
; APPLICANT: Larosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 100
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-840-459-100

Query Match          19.5%; Score 24; DB 10; Length 148;
Best Local Similarity 100.0%; Pred. No. 1e-14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 11  
US-09-809-739-22  
; Sequence 22, Application US/09809739  
; Patent No. US20020106369A1

; GENERAL INFORMATION:  
 ; APPLICANT: LaRosa, Gregory J.  
 ; APPLICANT: Horvath, Christopher  
 ; APPLICANT: Newman, Walter  
 ; APPLICANT: Jones, S. Tarran  
 ; APPLICANT: O'Brien, Siobhan H.  
 ; APPLICANT: O'Keefe, Theresa

RESULT 14  
US-09-840-459-20  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE OF INVENTION: METHODS OF USE THEREFOR  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanized sequence  
US-09-840-459-19

Query Match 17.1%; Score 21; DB 10; Length 117;  
Best Local Similarity 100.0%; Pred. No. 3.9e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 YAMNVRQAPGKGLEWVARIR 52  
DB 32 YAMNVRQAPGKGLEWVARIR 52  
|||||

RESULT 14  
US-09-840-459-20  
; Sequence 20, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: Larosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Slobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE OF INVENTION: METHODS OF USE THEREFOR  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanized sequence  
US-09-840-459-20

Query Match 17.1%; Score 21; DB 10; Length 117;  
Best Local Similarity 100.0%; Pred. No. 3.9e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 YAMNVRQAPGKGLEWVARIR 52  
DB 32 YAMNVRQAPGKGLEWVARIR 52  
|||||

RESULT 15  
US-09-828-708-12  
; Sequence 12, Application US/09828708  
; Patent No. US20020146753A1  
; GENERAL INFORMATION:  
; APPLICANT: Ditzel, H.  
; APPLICANT: Burton, D.  
; APPLICANT: Schaller, M.  
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their part  
; FILE OF INVENTION: autoimmune disease  
; FILE REFERENCE: 1361.005US1  
; CURRENT APPLICATION NUMBER: US/09/828,708  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-828-708-12

Query Match 13.8%; Score 17; DB 10; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 YAMNVRQAPGKGLEWV 48  
DB 25 YAMNVRQAPGKGLEWV 41  
|||||

Search completed: November 27, 2002, 07:32:53  
Job time : 12.3977 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 04:43:17 ; Search time 1253.48 Seconds  
(without alignments)  
7383.171 Million cell updates/sec

Title: US-09-893-615-88

Perfect score: 318

Sequence: 1 CAAATTCTCTCTCCAGTC.....GGACCATGCTGGAATAAGA 318

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_lov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	199	62.6	318	10	AF178619	AF178619 Mus muscu
2	172	54.1	318	6	AR096196	AR096196 Sequence
3	172	54.1	318	6	AR210595	AR210595 Sequence
4	172	54.1	321	6	AR096128	AR096128 Sequence
5	172	54.1	321	6	AR210527	AR210527 Sequence
6	172	54.1	335	6	AR096195	AR096195 Sequence
7	172	54.1	335	6	AR210594	AR210594 Sequence
8	161	50.6	318	10	AF178620	AF178620 Mus muscu
9	147	46.2	767	10	MMU231219	AF231219 Mus muscu
10	146	45.9	327	10	AF087025	AF087025 Mus muscu
11	142	44.7	300	10	AF137624	AF137624 Mus muscu
12	142	44.7	300	10	AF137625	AF137625 Mus muscu
13	142	44.7	312	10	S48339	S48339 Ig V kappa
14	141	44.3	302	10	MUSIGKAF	M4153 Mus musculu
15	140	44.0	1370	10	MUSIGKAAR	K01641 Mouse Ig ka
16	139	43.7	272	10	AF139247	AF139247 Mus muscu
17	138	43.4	300	10	MUSIGKAFI	M4157 Mus musculu
18	135	42.5	293	10	AF206030	AF206030 Mus muscu
19	135	42.5	360	10	AF029237	AF029237 Mus muscu
20	134	42.1	321	10	MMIG013	X58586 Mouse hybr
21	131	41.2	318	6	AX256247	AX256247 Sequence
22	131	41.2	318	12	AF277092	AF277092 Synthetic
23	131	41.2	765	6	AX057984	AX057984 Sequence
24	131	41.2	1239	6	AX057945	AX057945 Sequence
25	131	41.2	1280	6	AX057947	AX057947 Sequence
26	125	39.3	403	6	I09200	I09200 Sequence 40
27	125	39.3	403	10	MUSIGKCNA	MI7954 Mouse Ig re
28	124	39.0	307	10	AF139231	AF139231 Mus muscu
29	120	37.7	386	6	A20585	A20585 A5B7 human
30	120	37.7	705	6	A83197	A83197 Sequence 17
31	120	37.7	705	6	A8182927	A8182927 Sequence
32	120	37.7	732	6	A51863	A51863 Sequence 27
33	120	37.7	732	6	A83237	A83237 Sequence 57
34	120	37.7	732	6	AR085831	AR085831 Sequence
35	120	37.7	732	6	AR182961	AR182961 Sequence
36	120	37.7	1590	6	AX000421	AX000421 Sequence
37	120	37.7	1590	6	AX000422	AX000422 Sequence
38	120	37.7	3217	6	A83232	A83232 Sequence 52
39	120	37.7	3217	6	AR182956	AR182956 Sequence
40	117	36.8	403	6	I08812	I08812 Sequence 14
41	113	35.5	350	10	AY058906	AY058906 Mus muscu
42	109	34.3	264	10	MMU37887	U37887 Mus musculu
43	109	34.3	309	10	MMIGVKP4	X06517 Mouse Ig ac
44	107	33.6	435	6	A17967	A17967 Variable re
45	98	30.8	741	12	AY049714	AY049714 Synthetic

ALIGNMENTS

RESULT 1  
AF178619

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AF178619 318 bp mRNA linear ROD 22-MAY-2000  
Mus musculus 8-3 immunoglobulin light chain variable region mRNA,  
partial cds.

AF178619

AF178619.1 GI:5853221

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 318)

Puterman,C., Deocharan,B. and Diamond,B.

Molecular analysis of the autoantibody response in peptide-induced

autoimmunity  
J. Immunol. 164 (5), 2542-2549 (2000)  
MEDLINE  
20143847  
PUBMED  
10679092  
REFERENCE  
2 (bases 1 to 318)  
Putterman, C., Deocharan, B. and Diamond, B.  
Direct Submission  
TITLE  
Submitted (17-AUG-1999) Rheumatology/Medicine, Albert Einstein  
College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA  
FEATURES  
Location/Qualifiers  
1..318  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="8-3"  
/cell\_type="hybridoma"  
/note="from peptide-immunized mouse"  
<1..>318  
CDS  
/note="IGM kappa; anti-peptide antibody"  
/codon\_start=1  
/product="immunoglobulin light chain variable region"  
/protein\_id="A054372.1"  
/db\_xref="GI:5953222"  
/translation="KIVISQSPAILASPGQKVTWTCRASSSVYMLWYQQKPGSSPK  
PWISATSNLASGVPARFSGSGSTSYLSIRVEADAATYYCQOWSSNPPTFGGTR  
LEIK"  
BASE COUNT 76 a 92 c 78 g 72 t  
ORIGIN  
Query Match 62.6%; Score 199; DB 10; Length 318;  
Best Local Similarity 100.0%; Pred. No. 1.5e-110;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 106 CAGCAGAACGAGGATCCTCCCAACCTGGATTCTGCCACATCCAACTGGCTTCT 165  
DB 106 CACGAGAACGAGATCTCTCCCAACCTGGATTCTGCCACATCCAACTGGCTTCT 165  
QY 166 GGAGTCCCTGCTCGTTCACTGGCAGTGGTCTGGGACCTCTTACTCTTCACAATCAGC 225  
DB 166 GGAGTCCCTGCTCGTTCACTGGCAGTGGTCTGGGACCTCTTACTCTTCACAATCAGC 225  
QY 226 AGAGTGGAGCTGAAGATGCTGCACCTATTACTGCCAGCAGTGGAGTAGTACCCACC 285  
DB 226 AGAGTGGAGCTGAAGATGCTGCACCTATTACTGCCAGCAGTGGAGTAGTACCCACC 285  
QY 286 ACCTTCGGAGGGGGACCA 304  
DB 286 ACCTTCGGAGGGGGACCA 304  
RESULT 2  
AR096196  
LOCUS 318 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 104 from patent US 6005091.  
ACCESSION AR096196  
VERSION AR096196.1 GI:10024781  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 318)  
AUTHORS Blackburn, M. Neal., Church, W. Robert., Feuerstein, G. Zeev., Gross, M. Stuart., Nichols, A. John., Padlan, E. Augustin., Patel, A. Haribhai., and Sylvester, D. Robert.  
TITLE Nucleic acids encoding immunoglobulin domains  
JOURNAL Patent: US 6005091-A 104 21-DEC-1999;  
FEATURES Location/Qualifiers  
1..318  
/organism="unknown"  
BASE COUNT 79 a 91 c 78 g 70 t  
ORIGIN  
Query Match 54.1%; Score 172; DB 6; Length 318;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 52 AAGTCAACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACCTCTACCCAGCAG 111  
DB 52 AAGTCAACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACCTCTACCCAGCAG 111  
QY 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
DB 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
QY 172 CCTGCTCGCTTCACTGGCAGTGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
DB 172 CCTGCTCGCTTCACTGGCAGTGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
RESULT 3  
AR210595  
LOCUS 318 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 104 from patent US 6391299.  
ACCESSION AR210595  
VERSION AR210595.1 GI:21513363  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 318)  
AUTHORS Blackburn, M. Neal., Church, W. Robert., Feuerstein, G. Zeev., Gross, M. Stuart., Nichols, A. John., Padlan, E. Augustin., Patel, A. Haribhai., and Sylvester, D. Robert.  
TITLE Anti-factor IX/IXa antibodies  
JOURNAL Patent: US 6391299-A 104 21-MAY-2002;  
FEATURES Location/Qualifiers  
1..318  
/organism="unknown"  
BASE COUNT 79 a 91 c 78 g 70 t  
ORIGIN  
Query Match 54.1%; Score 172; DB 6; Length 318;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 52 AAGTCAACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACCTCTACCCAGCAG 111  
DB 52 AAGTCAACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACCTCTACCCAGCAG 111  
QY 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
DB 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
QY 172 CCTGCTCGCTTCACTGGCAGTGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
DB 172 CCTGCTCGCTTCACTGGCAGTGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
RESULT 4  
AR096128  
LOCUS 321 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 6 from patent US 6005091.  
ACCESSION AR096128  
VERSION AR096128.1 GI:10024649  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 318)  
AUTHORS Blackburn, M. Neal., Church, W. Robert., Feuerstein, G. Zeev., Gross, M. Stuart., Nichols, A. John., Padlan, E. Augustin., Patel, A. Haribhai., and Sylvester, D. Robert.  
TITLE Nucleic acids encoding immunoglobulin domains  
JOURNAL Patent: US 6005091-A 104 21-DEC-1999;  
FEATURES Location/Qualifiers  
1..318  
/organism="unknown"  
BASE COUNT 79 a 91 c 78 g 70 t  
ORIGIN  
Query Match 54.1%; Score 172; DB 6; Length 318;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

REFERENCE 1 (bases 1 to 321)  
AUTHORS Blackburn,M.Neal., Church,W.Robert., Feuerstein,G.Zeev., Gross,M.Stuart., Nichols,A.John., Padlan,E.Agustin., Patel,A.Haribhai., and Sylvester,D.Robert.  
TITLE Nucleic acids encoding immunoglobulin domains  
JOURNAL Patent: US 6005091-A 6 21-DEC-1999;  
FEATURES Location/Qualifiers  
source 1..321  
BASE COUNT 78 a 92 c 79 g 72 t  
ORIGIN

Query Match 54.1%; Score 172; DB 6; Length 321;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
Db 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
QY 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

RESULT 5  
AR210527  
LOCUS AR210527 321 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 6 from patent US 6391299.  
ACCESSION AR210527  
VERSION AR210527.1 GI:21513278  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 321)  
AUTHORS Blackburn,M.Neal., Church,W.Robert., Feuerstein,G.Zeev., Gross,M.Stuart., Nichols,A.John., Padlan,E.Agustin., Patel,A.Haribhai., and Sylvester,D.Robert.  
TITLE Anti-factor IX/IXa antibodies  
JOURNAL Patent: US 6391299-A 6 21-MAY-2002;  
FEATURES Location/Qualifiers  
source 1..321  
BASE COUNT 78 a 92 c 79 g 72 t  
ORIGIN

Query Match 54.1%; Score 172; DB 6; Length 321;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
Db 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
QY 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

REFERENCE 1 (bases 1 to 321)  
AUTHORS Blackburn,M.Neal., Church,W.Robert., Feuerstein,G.Zeev., Gross,M.Stuart., Nichols,A.John., Padlan,E.Agustin., Patel,A.Haribhai., and Sylvester,D.Robert.  
TITLE Anti-factor IX/IXa antibodies  
JOURNAL Patent: US 6391299-A 6 21-MAY-2002;  
FEATURES Location/Qualifiers  
source 1..321  
BASE COUNT 78 a 92 c 79 g 72 t  
ORIGIN

Query Match 54.1%; Score 172; DB 6; Length 321;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
Db 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
QY 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

RESULT 6  
AR096195  
LOCUS AR096195 335 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 102 from patent US 6005091.  
ACCESSION AR096195  
VERSION AR096195.1 GI:10024780  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 335)  
AUTHORS Blackburn,M.Neal., Church,W.Robert., Feuerstein,G.Zeev., Gross,M.Stuart., Nichols,A.John., Padlan,E.Agustin., Patel,A.Haribhai., and Sylvester,D.Robert.  
TITLE Nucleic acids encoding immunoglobulin domains  
JOURNAL Patent: US 6005091-A 102 21-DEC-1999;  
FEATURES Location/Qualifiers  
source 1..335  
BASE COUNT 80 a 97 c 85 g 73 t  
ORIGIN

Query Match 54.1%; Score 172; DB 6; Length 335;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
Db 112 AAGCCAGGATCCCTCCGCCAAACCTGGATTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171  
|||||  
QY 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

RESULT 7  
AR210594  
LOCUS AR210594 335 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 102 from patent US 6391299.  
ACCESSION AR210594  
VERSION AR210594.1 GI:21513362  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 335)  
AUTHORS Blackburn,M.Neal., Church,W.Robert., Feuerstein,G.Zeev., Gross,M.Stuart., Nichols,A.John., Padlan,E.Agustin., Patel,A.Haribhai., and Sylvester,D.Robert.  
TITLE Anti-factor IX/IXa antibodies  
JOURNAL Patent: US 6391299-A 102 21-MAY-2002;  
FEATURES Location/Qualifiers  
source 1..335  
BASE COUNT 80 a 97 c 85 g 73 t  
ORIGIN

Query Match 54.1%; Score 172; DB 6; Length 335;  
Best Local Similarity 99.6%; Pred. No. 6.2e-94;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCAGCTGGTACCAGCAG 111  
|||||



```
/gene="am4"
/Note="nonamer recombination signal"
BASE COUNT      210 a      180 c      149 g      228 t
ORIGIN
Query Match      46.2%; Score 147; DB 10; Length 767;
Best Local Similarity 99.5%; Pred. No. 1.4e-78;
Matches 197; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 90 TTACATGCACTGGTACCAAGCAGAGGATCCCTCCCAAAACCTGGATTCTGCCAC 149
|||||
DB 508 TTACATGCACTGGTACCAAGCAGAGGATCCCTCCCAAAACCTGGATTCTGCCAC 567
|||||

QY 150 ATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 209
|||||
DB 568 ATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 627
|||||

QY 210 CTCTCTCAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
|||||
DB 628 CTCTCTCAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 687
|||||

QY 270 GAGTAGTAACCCACCAC 287
|||||
DB 688 GAGTAGTAACCCACCAC 705
|||||

RESULT 10
LOCUS AF087025 327 bp DNA linear ROD 01-OCT-1998
DEFINITION Mus musculus hybrid 2-2H11 immunoglobulin light chain gene, partial
sequence.
ACCESSION AF087025
VERSION AF087025.1 GI:3676751
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 327)
AUTHORS Retter,M.W. and Nemazee,D.
TITLE Receptor editing occurs frequently during normal B cell development
JOURNAL J. Exp. Med. (1998) In press
REFERENCE 2 (bases 1 to 327)
AUTHORS Retter,M.W. and Nemazee,D.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-1998) Pediatrics K1023, National Jewish Medical
and Research Center, 1400 Jackson Street, Denver, CO 80206, USA
FEATURES
Location/Qualifiers
1..327
/organism="Mus musculus"
/strain="B10.D2 x Jkappa deficient)f1"
/db_xref="taxon:10090"
/rearranged
/Note="hybrid 2-2H11"
<1..327
misc_feature
/Note="productively rearranged immunoglobulin light chain
silenced by RS-IRSI rearrangement; VK4/5, JK5"
BASE COUNT      75 a      95 c      84 g      73 t
ORIGIN
Query Match      45.9%; Score 146; DB 10; Length 327;
Best Local Similarity 99.5%; Pred. No. 6.2e-78;
Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 90 TTACATGCACTGGTACCAAGCAGAGGATCCCTCCCAAAACCTGGATTCTGCCAC 149
|||||
DB 90 TTACATGCACTGGTACCAAGCAGAGGATCCCTCCCAAAACCTGGATTCTGCCAC 149
|||||

QY 150 ATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 209
|||||
DB 150 ATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 209
|||||

QY 210 CTCTCTCAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
|||||
DB 210 CTCTCTCAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
|||||

QY 270 GAGTAGTAACCCACCAC 282
|||||
DB 252 GAGTAGTAACCCAC 264
|||||

RESULT 12
LOCUS AF137625 300 bp mRNA linear ROD 03-JUN-1999
DEFINITION Mus musculus clone HV3PH4E2 anti-fluorescein immunoglobulin light
chain mRNA, partial cds.
ACCESSION AF137624
VERSION AF137624.1 GI:4972854
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 300)
AUTHORS van der Keyl,H., Gellad,Z.F. and Owen,J.A.
TITLE Disparity in the kinetics of onset of hypermutation in
immunoglobulin heavy and light chains
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 300)
AUTHORS van der Keyl,H., Gellad,Z.F. and Owen,J.A.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-1999) Biology, Haverford College, 370 Lancaster
Avenue, Haverford, PA 19041-1392, USA
FEATURES
Location/Qualifiers
1..300
/organism="Mus musculus"
/strain="Balb/c"
/db_xref="taxon:10090"
/clone="HV3P3G10"
/cell_type="hybridoma"
<1..300
/Note="from primary IgM antibody"
/codon_start=1
/product="anti-fluorescein immunoglobulin light chain"
/protein_id="AAD34819.1"
/db_xref="GI:4972855"
/translation="SPAILASPGKEKVTMTCRASSVSVMHWYQKPGSSPKPIYAT
SNLASEGVPARFSGSGSTSYSLTISRVEADDAITYCOOWSSNPWTFGGTKLEIK"
BASE COUNT      73 a      84 c      77 g      66 t
ORIGIN
Query Match      44.7%; Score 142; DB 10; Length 300;
Best Local Similarity 99.5%; Pred. No. 1.8e-75;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 90 TTACATGCACTGGTACCAAGCAGAGGATCCCTCCCAAAACCTGGATTCTGCCAC 149
|||||
DB 72 TTACATGCACTGGTACCAAGCAGAGGATCCCTCCCAAAACCTGGATTCTGCCAC 131
|||||

QY 150 ATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 209
|||||
DB 132 ATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 191
|||||

QY 210 CTCTCTCAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
|||||
DB 192 CTCTCTCAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 251
|||||

QY 270 GAGTAGTAACCCAC 282
|||||
DB 252 GAGTAGTAACCCAC 264
|||||

RESULT 12
LOCUS AF137625 300 bp mRNA linear ROD 03-JUN-1999
DEFINITION Mus musculus clone HV3PH4E2 anti-fluorescein immunoglobulin light
```

```
chain mRNA, partial cds.
ACCESSION AF137625
VERSION AF137625.1 GI:4972856
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 300)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL van der Keyl, H., Gellad, Z.F. and Owen, J.A.
AUTHORS Disparity in the kinetics of onset of hypermutation in
TITLE immunoglobulin heavy and light chains
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 300)
AUTHORS van der Keyl, H., Gellad, Z.F. and Owen, J.A.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-1999) Biology, Haverford College, 370 Lancaster
TITLE Avenue, Haverford, PA 19041-1392, USA
FEATURES
source Location/Qualifiers
1..300
/organism="Mus musculus"
/strain="Balb/c"
/db_xref="taxon:10090"
/clone="HV3PH4E2"
/cell_type="hybridoma"
<1..>300
/notes="from primary IgM antibody"
/codon_start=1
/product="anti-fluorescein immunoglobulin light chain"
/protein_id="AAD34820.1"
/db_xref="GI:4972857"
/translation="QSPAILASPGKEKVTMTCRASSSVSYMHVYQKPGSSPKWTYA
TSNLASGVPAFSGSGTYSYLISRVEADANTYTCQWSSNPWTFGGTKLEI"
BASE COUNT 71 a 85 c 78 g 66 t
ORIGIN
Query Match 44.7%; Score 142; DB 10; Length 300;
Best Local Similarity 99.5%; Pred. No. 1.8e-75;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 90 TTACATGCACTGGTGTACAGCAGAGCCAGGATCTCTCCCAAAACCCCTGGATTCTGCCAC 149
Db 75 TTACATGCACTGGTGTACAGCAGAGCCAGGATCTCTCCCAAAACCCCTGGATTCTGCCAC 134
QY 150 ATCCAACTGGCTTCTGGAGTCCCTGCTGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 209
Db 135 ATCCAACTGGCTTCTGGAGTCCCTGCTGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 194
QY 210 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
Db 195 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 254
QY 270 GAGTAGTAACCCA 282
Db 255 GAGTAGTAACCCA 267
RESULT 13
S48339
LOCUS S48339 312 bp mRNA linear ROD 08-MAY-1993
DEFINITION Ig V kappa -anti-Idiotypic Fab [mice, mRNA Partial, 312 nt].
ACCESSION S48339
VERSION S48339.1 GI:257749
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 312)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Kasai, Y., Herlyn, D., Sperlagh, M., Maruyama, H., Matsushita, S. and
Linnenbach, A.J.
TITLE Molecular cloning of murine monoclonal anti-Idiotypic Fab
JOURNAL J. Immunol. Methods 155 (1), 77-89 (1992)
```

```
93017981
MEDLINE 1383347
PUBMED GenBank staff at the National Library of Medicine created this
REMARK entry [NCBI gibbsq 118382] from the original journal article.
This sequence comes from Fig. 5.
FEATURES
source Location/Qualifiers
1..312
/organism="Mus sp."
/db_xref="taxon:10095"
1..312
/partial
/gene="Ig V-kappa"
/notes="anti-Idiotypic Fab"
1..312
/partial
/gene="Ig V-kappa"
/notes="anti-Idiotypic Fab; This sequence comes from Fig.
5"
/codon_start=1
/protein_id="AAB23708.1"
/db_xref="GI:257750"
/translation="VMTQSPAILASPGKEKVTMTCRASSSVSYMHVYQKPGSSPKPW
IVATSNLASGVPAFSGSGTYSYLISRVEADANTYTCQWSSNPWTFGGTKLEI"
BASE COUNT 76 a 88 c 80 g 68 t
ORIGIN
Query Match 44.7%; Score 142; DB 10; Length 312;
Best Local Similarity 99.5%; Pred. No. 1.8e-75;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 90 TTACATGCACTGGTGTACAGCAGAGCCAGGATCTCTCCCAAAACCCCTGGATTCTGCCAC 149
Db 84 TTACATGCACTGGTGTACAGCAGAGCCAGGATCTCTCCCAAAACCCCTGGATTCTGCCAC 143
QY 150 ATCCAACTGGCTTCTGGAGTCCCTGCTGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 209
Db 144 ATCCAACTGGCTTCTGGAGTCCCTGCTGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 203
QY 210 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
Db 204 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 263
QY 270 GAGTAGTAACCCA 282
Db 264 GAGTAGTAACCCA 276
RESULT 14
MUSIGKAPE
LOCUS Mus musculus Mouse Ig active kappa-chain mRNA V-region.
DEFINITION M64153
ACCESSION M64153.1 GI:196690
VERSION M64153.1
KEYWORDS V-region; Immunoglobulin light chain; immunoglobulin-kappa;
processed gene.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 302)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Caton, A.J., Stark, S.E., Kavalier, J., Staudt, L.M., Schwartz, D. and
Gerhard, W.
TITLE Many variable region genes are utilized in the antibody response of
JOURNAL BALB/c mice to the influenza virus A/PR/8/34 hemagglutinin
MEDLINE J. Immunol. 147 (5), 1675-1686 (1991)
PUBMED 91349593
COMMENT Specific for influenza virus A/PR/8/34 hemagglutinin.
FEATURES Location/Qualifiers
1..302
/organism="Mus musculus"
/strain="BALB/c"
```

```
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_line="H142-3H5"
/cell_type="B-cell hybridoma"
<1..3302
/codon_start=1
/product="immunoglobulin kappa-chain Vk-1"
/protein_id="AAB02375.1"
/db_xref="GI:196691"
/translation="QIVLSQSPAILSSPGEKVTMTCRASSSVSYMHVYQOKPGSSPK
PWIVATSNLASGVPARFSGSGTSLTISRVEADAATYYCQWSSNPYTFGGGT"
BASE COUNT      69 a      87 c      77 g      69 t
ORIGIN
Query Match      44.3%; Score 141; DB 10; Length 302;
Best Local Similarity 99.5%; Pred. No. 7.5e-75;
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 90 TTACATGCACCTGGTACACGAGAGCCAGGATCCTCCCCCAACCCCTGGATTTCGCCAC 149
Db 90 TTACATGCACCTGGTACACGAGAGCCAGGATCCTCCCCCAACCCCTGGATTTCGCCAC 149
Qy 150 ATCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 209
Db 150 ATCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTGGACCTCTTA 209
Qy 210 CTCCTCACAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
Db 210 CTCCTCACAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
Qy 270 GAGTAGTAACCC 281
Db 270 GAGTAGTAACCC 281

RESULT 15
MUSIGKAAR      1370 bp      DNA      linear      ROD 27-APR-1993
LOCUS          Mouse Ig kappa active V-region from 702/3 cells.
DEFINITION     K01641
ACCESSION      K01641
VERSION        K01641.1 GI:196460
KEYWORDS       C-region; V-region; immunoglobulin light chain;
               immunoglobulin-kappa.
SOURCE          Mouse 702/3 cell DNA.
ORGANISM       Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1 (bases 1 to 1370)
AUTHORS        Parslow,T.G., Blair,D.L., Murphy,W.J. and Granner,D.K.
TITLE          Structure of the 5' ends of immunoglobulin genes: a novel conserved
               sequence
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 81 (9), 2650-2654 (1984)
MEDLINE        84194064
PUBMED         6425835
COMMENT        Corrections to the sequence were sent by Dr. Richard E. Manroe and
               Dr. Tristram G. Parslow on Dec. 9, 1991.
FEATURES       source
               1..1370
               /organism="Mus musculus"
               /db_xref="taxon:10090"
               join(776..824,1003..1337)
               /partial
               /note="Ig kappa chain variable region"
               /codon_start=1
               /protein_id="AAA38693.1"
               /db_xref="GI:196461"
               /translation="MDFQVQIFSLISASVMSRGQIVLSQSPAILSSPGEKVTMT
               CRASSSVSYMHVYQOKPGSPKPIVATSNLASGVPARFSGSGTSLTISRVEAE
               DAATYYCQWSSNPRTFGGTTKLEIK"
               <776..824
               /note="Ig kappa chain variable region"
               /number=1
               825..1002

CDS
exon
intron
```

```
/note="kappa cds intron"
1003..>1337
/note="Ig kappa chain variable region"
/number=2
1338..>1369
/note="kappa cds intron J-C"
404 a      273 c      253 g      440 t
BASE COUNT
ORIGIN
Query Match      44.0%; Score 140; DB 10; Length 1370;
Best Local Similarity 100.0%; Pred. No. 2.9e-74;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 144 TGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGAC 203
Db 1163 TGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGAC 1222
Qy 204 CTCCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 263
Db 1223 CTCCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 1282
Qy 264 GCAGTGGAGTAGTAACCCAC 283
Db 1283 GCAGTGGAGTAGTAACCCAC 1302

Search completed: November 27, 2002, 06:25:23
Job time : 1264.48 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 27, 2002, 03:54:06 : Search time 170.803 Seconds  
(without alignments)  
4192.745 Million cell updates/sec

Title: US-09-893-615-88  
Perfect score: 318  
Sequence: 1 CAATGTGTTCTCCAGTC.....GGACCATCTGGAATAAGA 318

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	172	54.1	318	18	AAT79900	Anti-Factor IX MAB
2	172	54.1	318	24	ABK24005	Mouse-human light
3	172	54.1	321	18	AAT77377	Mouse anti-human F
4	172	54.1	321	24	ABK23937	Murine BC2 light c
5	172	54.1	335	18	AAT79899	Anti-Factor IX MAB
6	172	54.1	335	24	ABK24004	Murine BC2 light c
7	141	44.3	449	21	AAA43472	Mouse secreted exp
8	131	41.2	318	24	AA597128	Anti-NKG2D hybrido
9	131	41.2	765	22	AAC86590	DNA encoding a fus

10	131	41.2	1239	22	AAC86563	DNA encoding a fus
11	131	41.2	1280	22	AAC86564	Anti-CD20 single c
12	130	40.9	1925	21	AAA15019	DNA encoding a CD-
13	125	39.3	403	8	AAAT70972	2H7 VL sequence in
14	125	39.3	403	10	AAAT70972	2H7 Vh sequence.
15	125	39.3	404	20	AAV82358	Mouse antibody 2H7
16	125	39.3	404	22	AAH22071	2H7 light chain va
17	125	39.3	426	18	AAT36317	2H7 antibody light
18	125	39.3	426	18	AAT70869	2H7 light chain va
19	125	39.3	426	18	AAT51043	Coding sequence fo
20	125	39.3	426	19	AAV03927	Mouse 2H7 antibody
21	125	39.3	426	19	AAV18558	Mouse 2H7 antibody
22	125	39.3	426	19	AAV18554	Mouse 2H7 antibody
23	120	37.7	387	13	AAQ20983	Encodes Variable r
24	120	37.7	387	13	AAQ27350	Encodes ASB7 antib
25	120	37.7	705	20	AAV72047	Plasmid pNG3/ASB7V
26	120	37.7	732	17	AAT42508	Murine ASB57 Light
27	120	37.7	732	20	AAV72081	Plasmid pBE14/ASB7
28	120	37.7	1590	20	AAV24806	ASB7 F(ab')2 codin
29	120	37.7	3217	20	AAV72076	IRES-based ASB7 ch
30	107	33.6	435	12	AAQ15115	IL-2 chimeric anti
31	107	33.6	435	14	AAQ36613	Anti-IL2R-alpha an
32	96	30.2	96	20	AAV05587	Anti-Staph (HAY) 9
33	94	29.6	765	22	AAC86591	DNA encoding a fus
34	90	28.3	292	11	AAQ06213	VK domain of antib
35	90	28.3	292	17	AAT63502	Monoclonal antibod
36	90	28.3	292	17	AAT36662	Monoclonal antibod
37	90	28.3	292	20	AAZ23370	Mouse antibody MAK
38	90	28.3	292	21	AAZ88541	Mouse monoclonal a
39	72	22.6	321	13	AAQ25668	Sequence of the an
40	71	22.3	306	11	AAQ06228	VK domain of antib
41	71	22.3	306	17	AAT63506	Monoclonal antibod
42	71	22.3	306	17	AAT36660	Monoclonal antibod
43	71	22.3	309	14	AAQ45948	MAB BW 2128 light
44	71	22.3	331	18	AAT73613	cDNA encoding ligh
45	71	22.3	384	17	AAT15539	Mouse 5C7.29 monoc

ALIGNMENTS

RESULT 1  
AAT79900  
ID AAT79900 standard; cDNA; 318 BP.  
XX  
AC AAT79900:  
XX  
DT 27-DEC-1997 (first entry)  
XX  
DE Anti-Factor IX MAB chimeric light chain cDNA.  
XX  
KW Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;  
KW chimeric antibody; antibody engineering; light chain; ss.  
XX  
OS Chimeric Mus musculus.  
OS Chimeric Homo sapiens.  
OS Chimeric synthetic.  
XX  
PN WO9726010-A1.  
XX  
PD 24-JUL-1997.  
XX  
PF 17-JAN-1997; 97WO-US00759.  
XX  
PR 24-OCT-1996; 96US-0029119.  
PR 17-JAN-1996; 96US-0010108.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.  
XX  
PI Blackburn MN, Church WR, Feuerstein GZ, Gross MS;  
PI Nichols AJ, Padlan EA, Patel AH, Sylvester DR;  
XX

```

DR WPI: 1997-385117/35.
XX P-PSDB; AAW24532.
PT Inhibiting thrombosis with self-limiting antibody to coagulation
PT factor - avoids uncontrolled bleeding by providing only partial
PT inhibition
XX
PS Example 7; Page 128; 150pp; English.
XX
CC This cDNA sequence encodes a mouse-human chimeric antibody
CC light chain (AAW24532) in which the variable region is derived
CC from mouse anti-human factor IX monoclonal antibody BC2 cDNA (see
CC AAW79899) and human sequences from the immunoglobulin Rf-TS3'CL
CC framework. It was obtained by PCR amplification (see AAT79897-98)
CC of BC2 cDNA and insertion of the PCR product into F9H2HC 1-3 cDNA
CC (see AAT77374). Claimed anti-Factor IX chimeric antibodies are
CC useful in the treatment of thrombosis.
XX
SQ Sequence 318 BP; 79 A; 91 C; 78 G; 70 T; 0 other;
Query Match 54.1%; Score 172; DB 18; Length 318;
Best Local Similarity 99.6%; Pred. No. 4.7e-75;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 52 AAGTCAATGACCTTCGAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
DB 52 AAGTCAATGACCTTCGAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
QY 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCGGAGTC 171
DB 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCGGAGTC 171
QY 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
DB 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
QY 232 GAGGCTGAAGATGCTGCACCTTATTACTGCCAGCAGTGGAGTA 274
DB 232 GAGGCTGAAGATGCTGCACCTTATTACTGCCAGCAGTGGAGTA 274
RESULT 2
ABK24005
ID ABK24005 standard; cDNA; 318 BP.
XX
AC ABK24005;
XX
DT 09-APR-2002 (first entry)
XX
DE Mouse-human light chain DNA.
XX
KW Human; mouse; BC2; animal post-thromboembolic induced ischaemia;
KW thrombolytic agent; anti-factor IX antibody; plasminogen activator; ss;
KW thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;
KW vasotropic; cardiant; PCR primer; anti-respiratory syncytial virus;
KW heavy chain variable region; light chain variable region.
XX
OS Chimeric - Mus sp.
OS Chimeric - Homo sapiens.
XX
PN WO200187339-A1.
XX
PD 22-NOV-2001.
XX
PF 05-OCT-2000; 2000WO-US27438.
XX
PR 15-MAY-2000; 2000US-0571434.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;
XX
DR WPI: 2002-082944/11.

```

```

DR P-PSDB; AAU81002.
XX
PT Treating post-thromboembolic induced ischaemia in an animal by
PT administering anti-factor IX antibody in combination with a plasminogen
PT activator
XX
PS Example 7; Page 154-155; 163pp; English.
XX
CC The invention relates to a method for treating an animal
CC post-thromboembolic induced ischaemia or reducing a required dose of a
CC thrombolytic agent in treatment of an animal post-thromboembolic induced
CC ischaemia, comprising administering an anti-factor IX antibody or its
CC fragment, optionally in combination with a plasminogen activator or
CC thrombolytic agent. The method is useful for treating
CC post-thromboembolic-induced ischaemia, for preventing thromboembolic
CC stroke in an animal, and for reducing a required dose of a thrombolytic
CC agent. Sequences ABK23932-ABK24009 represent DNA molecules encoding
CC antibodies and PCR primers used in the method of the invention.
XX
SQ Sequence 318 BP; 79 A; 91 C; 78 G; 70 T; 0 other;
Query Match 54.1%; Score 172; DB 24; Length 318;
Best Local Similarity 99.6%; Pred. No. 4.7e-75;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 52 AAGTCAATGACCTTCGAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
DB 52 AAGTCAATGACCTTCGAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
QY 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCGGAGTC 171
DB 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCGGAGTC 171
QY 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
DB 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
QY 232 GAGGCTGAAGATGCTGCACCTTATTACTGCCAGCAGTGGAGTA 274
DB 232 GAGGCTGAAGATGCTGCACCTTATTACTGCCAGCAGTGGAGTA 274
RESULT 3
AAT77377
ID AAT77377 standard; cDNA; 321 BP.
XX
AC AAT77377;
XX
DT 26-DEC-1997 (first entry)
XX
DE Mouse anti-human Factor IX antibody BC2 light chain cDNA.
XX
KW Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;
KW humanised antibody; antibody engineering; light chain; CDR;
KW complementarity determining region; myocardial infarction;
KW angina; atrial fibrillation; stroke; kidney damage;
KW pulmonary embolism; deep vein thrombosis; coronary angioplasty;
KW disseminated intravascular coagulation; artificial organ; sepsis;
KW shunt; prosthesis; ss.
XX
OS Mus musculus.
XX
PN WO9736010-A1.
XX
PD 24-JUL-1997.
XX
PF 17-JAN-1997; 97WO-US00759.
XX
PR 24-OCT-1996; 96US-0029119.
PR 17-JAN-1996; 96US-0010108.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
DR (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

```

XX Blackburn MN, Church WR, Feuerstein GZ, Gross MS;  
PI Nichols AJ, Padian EA, Patel AH, Sylvester DR;  
XX WPI; 1997-385117/35.  
DR P-PSDB; AAW24520.  
XX Inhibiting thrombosis with self-limiting antibody to coagulation  
PT factor - avoids uncontrolled bleeding by providing only partial  
PT inhibition  
XX Example 5; Page 64; 150pp; English.  
XX This cDNA sequence encodes the light chain variable region (see  
CC AAW24520) of mouse anti-human Factor IX monoclonal antibody BC2.  
CC Claimed humanised antibodies (see AAW24510-18) contain CDRs (see  
CC AAW24504-09) of BC2 heavy and light chains inserted into framework  
CC regions of selected human antibody sequences. They have self-  
CC limiting neutralising activity, and are useful as anticoagulant  
CC agents in treatment of thrombosis associated with myocardial  
CC infarction, unstable angina, atrial fibrillation, stroke, renal  
CC damage, pulmonary embolism, deep vein thrombosis, percutaneous  
CC transluminal coronary angioplasty, disseminated intravascular  
CC coagulation, sepsis, or artificial organs, shunts or prostheses  
CC (claimed). Also claimed are chimeric antibodies (see AAT79900), and  
CC Fab and Fab'2 fragments. The claimed antibodies do not cause  
CC uncontrolled bleeding (contrast heparin and warfarin) since they  
CC provide only partial inhibition of coagulation.  
XX SQ Sequence 321 BP; 78 A; 92 C; 79 G; 72 T; 0 other;  
Query Match 54.1%; Score 172; DB 18; Length 321;  
Best Local Similarity 99.6%; Pred. No. 4.7e-75;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACTGGTACCAGCAG 111  
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACTGGTACCAGCAG 111  
QY 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAAACCTGGCTTCTGGAGTC 171  
Db 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAAACCTGGCTTCTGGAGTC 171  
QY 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
RESULT 4  
ABK23937  
ID ABK23937 standard; cDNA; 321 BP.  
XX AC ABK23937;  
XX DT 09-APR-2002 (first entry)  
XX DE Murine BC2 light chain variable region cDNA.  
XX Human; mouse; BC2; animal post-thromboembolic induced ischaemia;  
KW Thrombolytic agent; anti-factor IX antibody; plasminogen activator; ss;  
KW Thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;  
KW vasotropic; cardiant; PCR primer; anti-respiratory syncytial virus;  
KW heavy chain variable region; light chain variable region.  
XX OS Mus sp.  
XX PN WO200187339-A1.  
XX PD 22-NOV-2001.

XX 05-OCT-2000; 2000WO-US27438.  
PF 15-MAY-2000; 2000US-0571434.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;  
XX WPI; 2002-082944/11.  
DR P-PSDB; AAU80976.  
XX Treating post-thromboembolic induced ischaemia in an animal by  
PT administering anti-factor IX antibody in combination with a plasminogen  
PT activator  
XX Example 5; Page 94; 163pp; English.  
XX The invention relates to a method for treating an animal  
CC post-thromboembolic induced ischaemia or reducing a required dose of a  
CC thrombolytic agent in treatment of an animal post-thromboembolic induced  
CC ischaemia, comprising administering an anti-factor IX activator or ics  
CC fragment, optionally in combination with a plasminogen activator or  
CC thrombolytic agent. The method is useful for treating  
CC post-thromboembolic-induced ischaemia, for preventing thromboembolic  
CC stroke in an animal, and for reducing a required dose of a thrombolytic  
CC agent. Sequences ABK23932-ABK24009 represent DNA molecules encoding  
CC antibodies and PCR primers used in the method of the invention.  
XX SQ Sequence 321 BP; 78 A; 92 C; 79 G; 72 T; 0 other;  
Query Match 54.1%; Score 172; DB 24; Length 321;  
Best Local Similarity 99.6%; Pred. No. 4.7e-75;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACTGGTACCAGCAG 111  
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACTGGTACCAGCAG 111  
QY 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAAACCTGGCTTCTGGAGTC 171  
Db 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAAACCTGGCTTCTGGAGTC 171  
QY 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
RESULT 5  
AAT79899  
ID AAT79899 standard; cDNA; 335 BP.  
XX AC AAT79899;  
XX DT 27-DEC-1997 (first entry)  
XX DE Anti-Factor IX MAb BC2 light chain PCR product.  
XX Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;  
KW chimeric antibody; antibody engineering; light chain; ss.  
XX OS Chimeric Mus musculus.  
XX OS Chimeric synthetic.  
XX PN WO9726010-A1.  
XX PD 24-JUL-1997.  
XX PD 17-JAN-1997; 97WO-US00759.

```
XX 24-OCT-1996; 96US-0029119.
PR 17-JAN-1996; 96US-0010108.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
XX Blackburn MN, Church WR, Feuerstein GZ, Gross MS;
PI Nichols AJ, Padlan EA, Patel AH, Sylvestre DR;
XX
XX WPI: 1997-385117/35.
DR P-PSDB: AAW24531.
XX
XX Inhibiting thrombosis with self-limiting antibody to coagulation
PT factor - avoids uncontrolled bleeding by providing only partial
PT inhibition
XX
XX Example 7; Page 126; 150pp; English.
XX
XX This cDNA sequence was obtained by PCR amplification (see AAT79897
CC and AAT79898) of the light chain variable region (see also AAT77377)
CC of mouse anti-human factor IX monoclonal antibody BC2 cDNA. The
CC amplification resulted in the addition of ScaI, NariI ends to the
CC VL region. The PCR product was ligated into ScaI, NariI-digested
CC F9ZHC 1-3 (see AAT77374) and digested with ScaI, NariI to produce a
CC mouse-human chimeric light chain F9CHLC (see AAT79900, AAW24532).
CC Claimed anti-Factor IX chimeric antibodies are useful in the
CC treatment of thrombosis.
XX
XX Sequence 335 BP; 80 A; 97 C; 85 G; 73 T; 0 other;
SQ
Query Match 54.1%; Score 172; DB 18; Length 335;
Best Local Similarity 99.6%; Pred. No. 4.6e-75;
Matches 229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
DB 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
QY 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171
DB 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171
QY 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
DB 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274
RESULT 6
ABK24004
ID ABK24004 standard; cDNA; 335 BP.
XX
XX AC ABK24004;
XX
XX 09-APR-2002 (first entry)
XX
XX Murine BC2 light chain modified variable region DNA.
XX
XX Human; mouse; BC2; animal post-thromboembolic induced ischaemia;
XX thrombolytic agent; anti-factor IX antibody; plasminogen activator; ss;
XX thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;
XX vasotrophic; cardiant; PCR primer; anti-respiratory syncytial virus;
XX heavy chain variable region; light chain variable region.
XX
XX Mus sp.
XX OS Synthetic.
XX
XX PN WO200187339-A1.
XX
```

```
PD 22-NOV-2001.
XX
XX 05-OCT-2000; 2000WO-US27438.
XX
XX 15-MAY-2000; 2000US-0571434.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;
XX
XX WPI: 2002-082944/11.
DR P-PSDB: AAU81001.
XX
XX Treating post-thromboembolic induced ischaemia in an animal by
PT administering anti-factor IX antibody in combination with a plasminogen
PT activator
XX
XX Example 7; Page 152-153; 163pp; English.
XX
XX The invention relates to a method for treating an animal
CC post-thromboembolic induced ischaemia or reducing a required dose of a
CC thrombolytic agent in treatment of an animal post-thromboembolic induced
CC ischaemia, comprising administering an anti-factor IX antibody or its
CC fragment, optionally in combination with a plasminogen activator or
CC thrombolytic agent. The method is useful for treating thromboembolic
CC post-thromboembolic-induced ischaemia, for preventing thromboembolic
CC stroke in an animal, and for reducing a required dose of a thrombolytic
CC agent. Sequences ABK23932-ABK24009 represent DNA molecules encoding
CC antibodies and PCR primers used in the method of the invention.
XX
XX Sequence 335 BP; 80 A; 97 C; 85 G; 73 T; 0 other;
SQ
Query Match 54.1%; Score 172; DB 24; Length 335;
Best Local Similarity 99.6%; Pred. No. 4.6e-75;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
DB 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
QY 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171
DB 112 AAGCCAGGATCTCTCCCAACCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171
QY 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
DB 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274
RESULT 7
AAA43472
ID AAA43472 standard; cDNA; 449 BP.
XX
XX AC AAA43472;
XX
XX 21-AUG-2000 (first entry)
XX
XX Mouse secreted expressed sequence tag SEQ ID NO:47.
XX
XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
XX expressed sequence tag; EST; probe; chemotactic; proliferative;
XX immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
XX antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;
XX antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy;
XX vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
```

KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
XX  
OS Mus musculus.  
XX  
PN WO200021991-A1.  
XX  
PD 20-APR-2000.  
XX  
XX 15-OCT-1999; 99WO-US24206.  
XX  
PF 15-OCT-1998; 98US-0104436.  
XX  
XX (GENY ) GENETICS INST INC.  
XX  
XX Jacobs K, McCoy JM, LaValle ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Bowman MR;  
XX  
XX WPI; 2000-317938/27.  
XX  
XX Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (sESTs), useful for treating various disorders  
PT such as autoimmune, infectious, and central nervous system disorders -  
XX  
XX Claim 1; Page 214; 803pp; English.  
XX  
XX AAA43426 to AAA45925 represent specifically claimed secreted expressed  
CC sequence tags (sESTs), isolated from human, mouse, chicken and rat  
CC tissue sources. The sESTs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
CC antiaslathmic; vulnary; antiulcer; osteopathic; neuroprotective;  
CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
CC anticonvulsant; and antidepressant. The sESTs can be used for gene  
CC therapy and in vaccines. The sESTs are useful as probes for the  
CC identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
CC in the exemplification of the present invention.  
XX  
SQ Sequence 449 BP; 110 A; 124 C; 107 G; 108 T; 0 other;  
Query Match 44.3%; Score 141; DB 21; Length 449;  
Best Local Similarity 99.5%; Pred. No. 9.4e-60;  
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 90 TTACATGCACTGGTACCAGCAGAGCCAGGATCTCCCCCAACCCCTGGATTTGCCCAC 149  
DB 179 TTACATGCACTGGTACCAGCAGAGCCAGGATCTCCCCCAACCCCTGGATTTGCCCAC 238  
QY 150 ATCCACACCTGGCTCTGGAGTCCCTCGCTTCAGTGGCAGTGGCTGGGACCTCTTA 209  
DB 239 ATCCACACCTGGCTCTGGAGTCCCTCGCTTCAGTGGCAGTGGCTGGGACCTCTTA 298  
QY 210 CTCTCTCAACATCAGCAGAGTGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269  
DB 299 CTCTCTCAACATCAGCAGAGTGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 358  
QY 270 GAGTAGTAGACC 281  
DB 359 GAGTAGTAGACC 370

RESULT 8  
AAS97128  
ID AAS97128 standard; cDNA; 318 BP.  
XX  
AC AAS97128;  
XX  
DT 26-FEB-2002 (first entry)  
XX  
DE Anti-NKG2D hybridoma 6H7E7 variable light chain DNA.  
XX  
KW Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;  
KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;  
KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;  
KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; ss;  
KW sarcoma; leukemia; lymphoma; virus; bacterium; fungus; protozoan; DAPI0;  
KW helminth; cytostatic; antimicrobial; immunomodulatory; lln2D10; 6H7E7;  
KW 8G7C10; 6E5A7; PCR primer.  
XX  
OS Homo sapiens.  
XX  
PN WO200171005-A2.  
XX  
PD 27-SEP-2001.  
XX  
XX 26-MAR-2001; 2001WO-EP03414.  
XX  
XX 24-MAR-2000; 2000EP-0106467.  
XX  
XX (KUFE/) KUFER P.  
PI Kufer P, Riethmuller G, Lutterbuese R, Borschert K, Kischel R;  
PI Mayer M, Hofmeister R;  
XX  
WPI; 2002-055119/07.  
P-PSDB; AAU72833.  
XX  
XX Multifunctional polypeptides comprising binding sites that specifically  
PT recognise extracellular groups of the NKG2D receptor complex and  
PT domains which function as receptors or ligands, useful for treating  
PT cancers and infectious diseases -  
PS Example 3; Fig 16; 114pp; English.  
XX  
XX The invention relates to a multifunctional polypeptide comprising a  
CC domain with a binding site that specifically recognises an extracellular  
CC group of the NKG2D receptor complex and a second domain which functions  
CC as a receptor or ligand. The polypeptide and its associated  
CC polynucleotide are used for the preparation of a pharmaceutical  
CC composition for the treatment of cancer, infections and/or autoimmune  
CC conditions. The cancer may be a tumour of the head and neck, stomach,  
CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,  
CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,  
CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.  
CC The infectious diseases can be caused by viruses, bacteria, fungi,  
CC protozoa or helminths. The autoimmune diseases include multiple  
CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior  
CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent  
CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and  
CC autoimmune hepatitis. The sequences represent the NKG2D receptor DNA, DNA  
CC encoding the polypeptides of the invention and PCR primers used to  
CC amplify the DNA sequences.

SQ Sequence 318 BP; 76 A; 92 C; 80 G; 70 T; 0 other;

Query Match 41.2%; Score 131; DB 24; Length 318;  
Best Local Similarity 99.5%; Pred. No. 8.2e-55;  
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 90 TTACATGCACTGGTACCAGCAGAGCCAGGATCTCCCCCAACCCCTGGATTTGCCCAC 149  
DB 90 TTACATGCACTGGTACCAGCAGAGCCAGGATCTCCCCCAACCCCTGGATTTGCCCAC 149

QY 150 ATCCAACCTGGCTTCTGAGTCCCTGCTTCAGTGGCAGTGGTCTGGGACCTCTTA 209  
|||||  
Db 150 ATCCAACCTGGCTTCTGAGTCCCTGCTTCAGTGGCAGTGGTCTGGGACCTCTTA 209  
|||||  
QY 210 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269  
|||||  
Db 210 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269  
|||||  
QY 270 GA 271  
||  
Db 270 GA 271  
||  
RESULT 9  
AAC86590  
ID AAC86590 standard; DNA; 765 BP.  
XX  
AC AAC86590;  
XX  
DT 02-APR-2001 (first entry)  
XX  
DE DNA encoding a fusion of a single chain antibody and streptavidin.  
XX  
KW Streptavidin; tumour cell; cancer; adenocarcinoma;  
KW hematological malignancy; ss.  
XX  
OS Synthetic.  
OS Streptomyces avidinii.  
OS Unidentified.  
XX  
PN WO200075333-A1.  
XX  
XX 14-DEC-2000.  
XX  
PF 05-JUN-2000; 2000WO-US15595.  
XX  
PR 07-JUN-1999; 99US-0137900.  
PR 03-DEC-1999; 99US-0168976.  
XX  
PA (NEOR-) NEORX CORP.  
XX  
XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
PI WPI; 2001-091213/10.  
XX  
DR New vector constructs for expressing genomic streptavidin fusion  
XX proteins which are useful for targeting tumour cells associated with  
PT cancer, e.g. adenocarcinomas -  
PT  
XX Example 5; Page 95; 100pp; English.  
PS  
XX The present sequence encodes a fusion of an anti-CD20 single chain  
CC antibody and streptavidin. The fusion protein is expressed using  
CC vectors of the invention. The specification describes vector constructs  
CC for expressing streptavidin fusion proteins. The vector comprises a  
CC nucleic acid encoding streptavidin or its functional variant operatively  
CC linked to a promoter, and a cloning site for insertion of a second  
CC nucleic acid sequence encoding a polypeptide to be fused with  
CC streptavidin, interposed between the promoter and the first nucleic  
CC acid sequence. Alternatively, the vector construct comprises a nucleic  
CC acid, operatively linked to a promoter, encoding a polypeptide to be  
CC fused with streptavidin, and a cloning site for insertion of a second  
CC nucleic acid encoding at least 129 amino acids of streptavidin or its  
CC functional variant. The fusion proteins are useful for targeting tumour  
CC cells, particularly tumour cells associated with cancer,  
CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
CC is useful for expressing of streptavidin fusion proteins. In particular,  
CC these are useful as tools for medical diagnostics and therapeutic  
CC purposes, e.g. for detecting the presence or absence of, or treating, a  
CC target site within a mammalian host.  
XX  
SQ Sequence 765 BP; 169 A; 201 C; 230 G; 165 T; 0 other;

Query Match 41.2%; Score 131; DB 22; Length 765;  
Best Local Similarity 99.5%; Pred. No. 8e-55;  
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 90 TTACATGCACCTGGTACCAGCAGAGAGCCAGGATCCTCCCCAAACCTTGGATTCTGCCAC 149  
|||||  
Db 537 TTACATGCACCTGGTACCAGCAGAGAGCCAGGATCCTCCCCAAACCTTGGATTCTGCCAC 596  
|||||  
QY 150 ATCCAACCTGGCTTCTGAGTCCCTGCTTCAGTGGCAGTGGTCTGGGACCTCTTA 209  
|||||  
Db 597 ATCCAACCTGGCTTCTGAGTCCCTGCTTCAGTGGCAGTGGTCTGGGACCTCTTA 656  
|||||  
QY 210 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269  
|||||  
Db 657 CTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 716  
|||||  
QY 270 GA 271  
||  
Db 717 GA 718  
||  
RESULT 10  
AAC86563  
ID AAC86563 standard; DNA; 1239 BP.  
XX  
AC AAC86563;  
XX  
DT 02-APR-2001 (first entry)  
XX  
DE DNA encoding a fusion of anti-CD20 single chain antibody/streptavidin.  
XX  
KW Streptavidin; tumour cell; cancer; adenocarcinoma;  
KW hematological malignancy; ss.  
XX  
OS Synthetic.  
OS Streptomyces avidinii.  
OS Homo sapiens.  
XX  
PN WO200075333-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 05-JUN-2000; 2000WO-US15595.  
XX  
PR 07-JUN-1999; 99US-0137900.  
PR 03-DEC-1999; 99US-0168976.  
XX  
PA (NEOR-) NEORX CORP.  
XX  
XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
PI WPI; 2001-091213/10.  
XX  
DR P-PSDB; AAB30694.  
XX  
PT New vector constructs for expressing genomic streptavidin fusion  
XX proteins which are useful for targeting tumour cells associated with  
PT cancer, e.g. adenocarcinomas -  
XX  
XX Example 2; Fig 11A; 100pp; English.  
PS  
XX The present sequence encodes a fusion of an anti-CD20 single chain  
CC antibody (B9E9) streptavidin. The fusion protein is expressed using  
CC vectors of the invention. The specification describes vector constructs  
CC for expressing streptavidin fusion proteins. The vector comprises a  
CC nucleic acid encoding streptavidin or its functional variant operatively  
CC linked to a promoter, and a cloning site for insertion of a second  
CC nucleic acid sequence encoding a polypeptide to be fused with  
CC streptavidin, interposed between the promoter and the first nucleic  
CC acid sequence. Alternatively, the vector construct comprises a nucleic  
CC acid, operatively linked to a promoter, encoding a polypeptide to be  
CC fused with streptavidin, and a cloning site for insertion of a second  
CC nucleic acid encoding at least 129 amino acids of streptavidin or its

CC functional variant. The fusion proteins are useful for targeting tumour  
CC cells, particularly tumour cells associated with cancer,  
CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
CC is useful for expressing of streptavidin fusion proteins. In particular,  
CC these are useful as tools for medical diagnostics and therapeutic  
CC purposes, e.g. for detecting the presence or absence of, or treating, a  
CC target site within a mammalian host.

XX Sequence 1239 BP; 270 A; 392 C; 356 G; 221 T; 0 other;

Query Match 41.2%; Score 131; DB 22; Length 1239;  
Best Local Similarity 99.5%; Pred. No. 7.8e-55;  
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 90 TTACATGCACCTGGTACACAGCAGAGCCAGGATCCCTCCCAACCCCTGGATTTCGCCAC 149

Db 90 TTACATGCACCTGGTACACAGCAGAGCCAGGATCCCTCCCAACCCCTGGATTTCGCCAC 149

Oy 150 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 209

Db 150 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 209

Oy 210 CTCCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCACAGTG 269

Db 210 CTCCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCACAGTG 269

Oy 270 GA 271

Db 270 GA 271

RESULT 11

ID AAC86564 standard; DNA; 1280 BP.

XX AAC86564;

XX 02-APR-2001 (first entry)

XX Anti-CD20 single chain antibody/streptavidin fusion protein cassette.

XX Streptavidin; tumour cell; cancer; adenocarcinoma;

XX hematological malignancy; ss.

XX Synthetic.

XX Streptomyces avidinii.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 3..1274

XX /\*tag= a

XX /product= "anti-CD20 scFv and streptavidin fusion"

XX WO200075333-A1.

XX 14-DEC-2000.

XX 05-JUN-2000; 2000WO-US15595.

XX 07-JUN-1999; 99US-0137900.

XX 03-DEC-1999; 99US-0168976.

XX (NEOR-) NEORX CORP.

XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;

XX WPI; 2001-091213/10.

XX P-PSDB; AAB30695.

XX New vector constructs for expressing genomic streptavidin fusion

XX proteins which are useful for targeting tumour cells associated with

XX cancer, e.g. adenocarcinomas .

PS Example 2; Fig 11C; 100pp; English.

XX The present sequence encodes a fusion of an anti-CD20 single chain  
XX antibody (B9P9) streptavidin. The fusion protein is expressed using  
XX vectors of the invention. The specification describes vector constructs  
XX for expressing streptavidin fusion proteins. The vector comprises a  
XX nucleic acid encoding streptavidin or its functional variant operatively  
XX linked to a promoter, and a cloning site for insertion of a second  
XX nucleic acid sequence encoding a polypeptide to be fused with  
XX streptavidin, interposed between the promoter and the first nucleic  
XX acid sequence. Alternatively, the vector construct comprises a nucleic  
XX acid, operatively linked to a promoter, encoding a polypeptide to be  
XX fused with streptavidin, and a cloning site for insertion of a second  
XX nucleic acid encoding at least 129 amino acids of streptavidin or its  
XX functional variant. The fusion proteins are useful for targeting tumour  
XX cells, particularly tumour cells associated with cancer, the vector construct  
XX e.g. adenocarcinomas or hematological malignancies. The vector construct  
XX is useful for expressing of streptavidin fusion proteins. In particular,  
XX these are useful as tools for medical diagnostics and therapeutic  
XX purposes, e.g. for detecting the presence or absence of, or treating, a  
XX target site within a mammalian host.

SQ Sequence 1280 BP; 267 A; 397 C; 388 G; 228 T; 0 other;

Query Match 41.2%; Score 131; DB 22; Length 1280;

Best Local Similarity 99.5%; Pred. No. 7.8e-55;

Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 90 TTACATGCACCTGGTACACAGCAGAGCCAGGATCCCTCCCAACCCCTGGATTTCGCCAC 149

Db 545 TTACATGCACCTGGTACACAGCAGAGCCAGGATCCCTCCCAACCCCTGGATTTCGCCAC 604

Oy 150 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 209

Db 605 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTA 664

Oy 210 CTCCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCACAGTG 269

Db 665 CTCCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCACAGTG 724

Oy 270 GA 271

Db 725 GA 726

RESULT 12

AAA15019

ID AAA15019 standard; DNA; 1925 BP.

XX AAA15019;

XX 21-AUG-2000 (first entry)

XX DNA encoding a CD-20 specific chimeric receptor.

XX CD20-specific receptor; CD-20 specific redirected T cell; leukemia;

XX CD20+ malignancy; non-Hodgkin's lymphoma; myeoablative chemotherapy;

XX stem cell rescue; autoimmune disease; lupus; rheumatoid arthritis; ss.

XX Synthetic.

XX Mus sp.

XX Key Location/Qualifiers

XX CDS 27..1928

XX /\*tag= a

XX WO200023573-A2.

XX 27-APR-2000.

XX 20-OCT-1999; 99WO-US24484.

XX 20-OCT-1998; 98US-0105014.





ID AAN91147 standard; DNA; 403 BP.  
XX AC AAN91147;  
XX DT 06-JUL-1990 (first entry)  
XX DE 2H7 Vh sequence.  
XX KW Antibodies; passive immunisation; pH3-6a; ss.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX CDS 20..403  
FT FT /\*tag= a  
FT FT misc\_feature 320..352 b  
FT FT /\*tag= b  
FT FT /note="JK5 region."  
XX PN W08900999-A.  
XX PN 01-NOV-1985;  
XX PD 09-FEB-1989.  
XX PF 25-JUL-1988; 88WO-US02514.  
XX PR 24-JUL-1987; 87US-0077528.  
XX PA (ITGE-) INT GENETIC ENG INC.  
XX PI Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;  
XX PT WPI; 1989-061144/08.  
XX DR P-PSDB; AAP94778.  
XX PT Polynucleotide(s) encoding Immunoglobulin molecules -  
PT used for efficient prodn. of chimeric human or non-human or  
PT class switched antibodies.  
XX PS Disclosure; ; 7pp; English.  
XX CC Sequence, derived from M13 subclones of gene fragments, carries the  
CC variable region of chimeric immunoglobulin sequence. The antibodies are  
CC useful in passive immunisation avoiding negative reactions. They are  
CC also useful in assaying and in vitro imaging.  
XX SQ Sequence 403 BP; 100 A; 112 C; 93 G; 98 T; 0 other;

Query Match 39.3%; Score 125; DB 10; Length 403;  
Best Local Similarity 100.0%; Pred. No. 7.5e-52;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 149 CATCCAACCTGGCTTCTCGAGTCCCTGCTTCAGTGGCAGTGGGCTGGGACCTCTT 208  
Db 234 CATCCAACCTGGCTTCTCGAGTCCCTGCTTCAGTGGCAGTGGGCTGGGACCTCTT 293  
Qy 209 ACTCTCTCACAAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGT 268  
Db 294 ACTCTCTCACAAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGT 353  
Qy 269 GGAGT 273  
Db 354 GGAGT 358

## RESULT 15

AAV82358

ID AAV82358 standard; DNA; 404 BP.

XX AC AAV82358;  
XX DT 30-MAR-1999 (first entry)

XX DE Mouse antibody 2H7 light chain variable region encoding sequence.

XX KW pelB pectate lyase; secretion signal; chimeric antibody;  
KW light chain; B-cell antigen; antibody 2H7; ss.  
XX OS Mus sp.  
XX FH Key Location/Qualifiers  
XX CDS 21..404  
FT FT /\*tag= a  
FT FT /note="partial CDS"  
XX PN US5846818-A.  
XX PD 08-DEC-1998.  
XX PF 06-JUN-1995; 95US-0472696.  
XX PR 29-MAR-1990; 90US-0501092.  
XX PR 01-NOV-1985; 85US-0793980.  
XX PR 24-JUL-1987; 87US-0077528.  
XX PR 11-JAN-1988; 88US-0142039.  
XX PR 08-DEC-1992; 92US-0987555.  
XX PR 22-FEB-1993; 93US-0020671.  
XX PR 09-DEC-1994; 94US-0357234.  
XX PR 06-JUN-1995; 95US-0472696.  
XX PA (XOMA ) XOMA CORP.  
XX PI Better M, Horowitz AH, Lei S, Liu AY, Robinson RR;  
PI Wall R, Wilcox GL;  
XX DR WPI; 1999-059072/05.  
XX DR P-PSDB; AAW89541.  
XX PT pelB pectate lyase signal sequence - and vector for expression of  
PT secreted proteins in Gram-negative bacteria  
XX PS Example 4; Fig 22; 98pp; English.  
XX CC The present sequence encodes the mouse antibody 2H7 light chain  
CC variable region. Antibody 2H7 is specific for human B-cell antigen. The  
CC antibody sequence was used to construct a chimeric human-mouse antibody,  
CC in the course of the invention. The chimeric antibody is expressed in a  
CC secretion vector comprising a pelB pectate lyase secretion signal  
CC peptide. The pelB pectate lyase secretion sequence is useful for  
CC producing a protein such as a chimeric antibody in a bacterial host.  
XX SQ Sequence 404 BP; 100 A; 113 C; 93 G; 98 T; 0 other;

Query Match 39.3%; Score 125; DB 20; Length 404;  
Best Local Similarity 100.0%; Pred. No. 7.5e-52;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 149 CATCCAACCTGGCTTCTCGAGTCCCTGCTTCAGTGGCAGTGGGCTGGGACCTCTT 208  
Db 235 CATCCAACCTGGCTTCTCGAGTCCCTGCTTCAGTGGCAGTGGGCTGGGACCTCTT 294  
Qy 209 ACTCTCTCACAAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGT 268  
Db 295 ACTCTCTCACAAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGT 354  
Qy 269 GGAGT 273  
Db 355 GGAGT 359

Search completed: November 27, 2002, 05:37:46  
Job time : 177.303 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 05:31:20 ; Search time 1316.44 Seconds  
(without alignments)  
3912.201 Million cell updates/sec

Title: US-09-893-615-88  
Perfect score: 318  
Sequence: 1 CAATTTCTCTCTCCAGTC.....GGACCATGCTGGAATAAGA 318

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_htc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_htc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_Other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_rod: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	28.9	639	13	BC968615 602836416
2	89	28.0	991	13	BC968852 602836142
3	81	25.5	413	10	BE624609 uu24c05.y
4	81	25.5	532	13	BI104783 602891329
5	81	25.5	569	13	BC964957 602829274
6	81	25.5	639	12	BF583310 602101781

7	71	22.3	396	9	AA691311	AA691311 vs14f01.r
8	71	22.3	960	12	BF581607	BF581607 602099949
9	70	22.0	918	14	BQ946795	BQ946795 AGENCOURT
10	68	21.4	959	14	BQ939046	BQ939046 AGENCOURT
11	67	21.1	798	13	BG968518	BG968518 602835104
12	64	20.1	755	13	BI150509	BI150509 602915167
13	63	19.8	950	12	BF138873	BF138873 601781783
14	60	18.9	280	12	BF149077	BF149077 uu88f09.y
15	59	18.6	918	14	BQ931133	BQ931133 AGENCOURT
16	53	16.7	728	13	BF015569	BF015569 uy23c09.y
17	53	16.7	728	13	BG962985	BG962985 602827980
18	53	16.7	865	13	BG969179	BG969179 602830945
19	53	16.7	878	12	BF137585	BF137585 601783111
20	53	16.7	934	14	BQ934947	BQ934947 AGENCOURT
21	52	16.4	711	12	BF142976	BF142976 601791755
22	52	16.4	862	13	BG962582	BG962582 602829965
23	52	16.4	865	13	BG964349	BG964349 602831966
24	52	16.4	953	12	BF583869	BF583869 602096867
25	51	16.0	857	14	BQ926839	BQ926839 AGENCOURT
26	50	15.7	472	10	BE289692	BE289692 601088922
27	49	15.4	257	12	BG148606	BG148606 uu84f12.y
28	49	15.4	312	12	BG090240	BG090240 ut58e10.y
29	49	15.4	766	13	BG967034	BG967034 602834270
30	49	15.4	1000	14	BQ937284	BQ937284 AGENCOURT
31	48	15.1	688	17	AZ395191	AZ395191 1M0164K24
32	48	15.1	793	13	BG965736	BG965736 602830592
33	48	15.1	807	17	BH035104	BH035104 RPCI-24-2
34	47	14.8	722	12	BF140170	BF140170 601786735
35	46	14.5	1058	13	BG964881	BG964881 602831773
36	45	14.2	721	17	AZ834202	AZ834202 2M0116K13
37	41	12.9	620	10	BE368918	BE368918 601221562
38	41	12.9	743	13	BI106381	BI106381 602892928
39	41	12.9	793	13	BI105963	BI105963 602890917
40	41	12.9	871	12	BF165585	BF165585 601777486
41	40	12.6	408	12	BF016419	BF016419 uy41a08.y
42	40	12.6	876	12	BF141750	BF141750 601787615
43	40	12.6	918	12	BF142385	BF142385 601786637
44	40	12.6	948	12	BF139176	BF139176 601783324
45	39	12.3	871	14	BQ956722	BQ956722 AGENCOURT

## ALIGNMENTS

RESULT 1	BC968615	602836416F1	NCI_CGAP_Co24	Mus musculus	cDNA clone	linear	EST 12-JUN-2001
LOCUS	BC968615	mRNA sequence.					
DEFINITION	BC968615	mRNA sequence.					
ACCESSION	BC968615	GI:14356252					
VERSION	BC968615.1	EST.					
KEYWORDS	EST.						
SOURCE	house mouse.						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.						
TITLE	1 (bases 1 to 639)						
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.						
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)						
	Unpublished (1999)						
	Contact: Robert Strausberg, Ph.D.						
	Email: cgapbs-re@mail.nih.gov						
	Tissue Procurement: Jeffrey E. Green, M.D.						
	CDNA Library Preparation: Life Technologies, Inc.						
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)						
	DNA Sequencing by: Incyte Genomics, Inc.						
	Clone distribution: MGC clone distribution information can be						
	found through the I.M.A.G.E. Consortium/LLNL at:						
	http://image.llnl.gov						
	Plate: LLAM11007 row: o column: 04						
	High quality sequence stop: 639.						
FEATURES	Location/Qualifiers						
source	1..639						

```
/organism="Mus musculus"
/db_strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      172 a   173 c   151 g   143 t
ORIGIN
Query Match      28.9%; Score 92; DB 13; Length 639;
Best Local Similarity 99.0%; Pred. No. 7.3e-36;
Matches 192; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  90 TTACATGCACCTGTACCAGCAGAGCCAGGATCCTCCCAACACCTGGATTCTGCCAC 149
Db  164 TTACATGCACCTGTACCAGCAGAGCCAGGATCCTCCCAACACCTGGATTCTGCCAC 223
QY  150 ATCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 209
Db  224 ATCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 283
QY  210 CTCTCTCAACATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTATTACTGCCAGCAGTG 269
Db  284 CTCTCTCAACATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTATTACTGCCAGCAGTG 343
QY  270 GAGTAGTAACCCAC 283
Db  344 GAGTAGTAACCCAC 357

RESULT 2
BG968852
LOCUS      991 bp      mRNA      linear      EST 12-JUN-2001
DEFINITION      602836142F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4990427 5',
mRNA sequence.
ACCESSION      BG968852
VERSION
KEYWORDS
SOURCE
ORGANISM      house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 991)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Library Arrayed by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1006 row: d column: 12
High quality sequence stop: 620.
FEATURES
source
1. .991
Location/Qualifiers
/organism="Mus musculus"
/db_strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="Soares_mammary_gland_NLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dt) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      106 a   113 c   94 g   100 t
ORIGIN
Query Match      25.5%; Score 81; DB 10; Length 413;
Best Local Similarity 100.0%; Pred. No. 2.6e-30;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  147 CACATCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTC 206
Db  243 CACATCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTC 302
QY  207 TTACTCTCTCACATCAGCAG 227
Db  303 TTACTCTCTCACATCAGCAG 323
```

```

RESULT 4
B1104783
LOCUS      602891329f1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5036622 5',
DEFINITION mRNA sequence.
ACCESSION B1104783
VERSION   B1104783.1 GI:14555676
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 532)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Gilbert Smith, Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L1AM1101 row: 1 column: 07
           High quality sequence stop: 532.
           Location/Qualifiers
             1..532
               /organism="Mus musculus"
               /strain="CZECH II"
               /db_xref="taxon:10090"
               /clone="IMAGE:5036622"
               /clone_lib="NCI_CGAP_Lu29"
               /tissue_type="spontaneous tumor, metastatic to mammary.
               Stem cell origin."
               /lab_host="DH10B"
               /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
               Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
               Library constructed by Life Technologies. Investigator
               providing samples: Gilbert Smith, NIH"
BASE COUNT 139 a 148 c 113 g 132 t
ORIGIN
Query Match      25.5%; Score 81; DB 13; Length 532;
Best Local Similarity 100.0%; Pred. No. 2.8e-30;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 147 CACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGTCTGGGACCTC 206
|||||
Db 248 CACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGTCTGGGACCTC 307
|||||

Oy 207 TTACTCTCTCAATCAGCAG 227
|||||
Db 308 TTACTCTCTCAATCAGCAG 328
|||||

RESULT 5
BG964957
LOCUS      602829274f1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4983883 5',
DEFINITION mRNA sequence.
ACCESSION BG964957
VERSION   BG964957.1 GI:14352594
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 569)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.

```

```

TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Jeffrey E. Green, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L1AM10989 row: c column: 20
           High quality sequence stop: 569.
           Location/Qualifiers
             1..569
               /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone="IMAGE:4983883"
               /clone_lib="NCI_CGAP_Co24"
               /lab_host="DH10B (T1 phage-resistant)"
               /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
               Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
               Average insert size 1.6 kb. Constructed by Life
               Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 144 a 155 c 132 g 138 t
ORIGIN
Query Match      25.5%; Score 81; DB 13; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.8e-30;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 147 CACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGTCTGGGACCTC 206
|||||
Db 211 CACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGTCTGGGACCTC 270
|||||

Oy 207 TTACTCTCTCAATCAGCAG 227
|||||
Db 271 TTACTCTCTCAATCAGCAG 291
|||||

RESULT 6
BF583310
LOCUS      602101781f1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4224692 5',
DEFINITION mRNA sequence.
ACCESSION BF583310
VERSION   BF583310.1 GI:11657028
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 639)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Jeffrey E. Green, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L1AM9815 row: b column: 21
           High quality sequence stop: 636.
           Location/Qualifiers
             1..639
               /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"

```

```

/clone="IMAGE:4224692"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/Note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      169 a 174 c 149 g 147 t
ORIGIN
Query Match      25.5%; Score 81; DB 12; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 CACATCAACCTGGCTTCCTGGAGTCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTC 206
|||||
Db 222 CACATCAACCTGGCTTCCTGGAGTCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTC 281
|||||

QY 207 TTACTCTCTCACAAATCAGCAG 227
|||||
Db 282 TTACTCTCTCACAAATCAGCAG 302
|||||

RESULT 7
AA691311
LOCUS
DEFINITION
vsl4f01.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
clone IMAGE:1138201 5' similar to gb:X67211 M.musculus rearranged
immunoglobulin kappa light chain (MOUSE);, mRNA sequence.
ACCESSION
AA691311
VERSION
AA691311.1 GI:2692247
KEYWORDS
EST.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 396)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, F., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:619473
Seq primer: -28ml3 rev2 ET from Amerham.
FEATURES
Location/Qualifiers
1..396
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1138201"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/Note="Vector: pTT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo(dT) primer
[5'TGCTTACGATCTGAAGTGGAGCGGCCCGCCCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGATCCTTGT], digested with Not I and cloned
into the Not I and Eco RI sites of the modified pTT73
vector. Library constructed by Bob Barstead. "
BASE COUNT      104 a 107 c 91 g 94 t
ORIGIN
Query Match      22.3%; Score 71; DB 9; Length 396;
Best Local Similarity 100.0%; Pred. No. 3.2e-25;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 CTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTC 216
|||||
Db 248 CTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTC 307
|||||

QY 217 ACAATCAGCAG 227
|||||
Db 308 ACAATCAGCAG 318
|||||

RESULT 8
BF581607
LOCUS
DEFINITION
602099949f1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4219669 5',
mRNA sequence.
ACCESSION
BF581607
VERSION
BF581607.1 GI:11655319
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 960)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
http://image.llnl.gov
Plate: LLAM9802 row: a column: 14
High quality sequence start: 3
High quality sequence stop: 631.
FEATURES
Location/Qualifiers
1..960
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4219669"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/Note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      282 a 259 c 221 g 198 t
ORIGIN
Query Match      22.3%; Score 71; DB 12; Length 960;
Best Local Similarity 100.0%; Pred. No. 3.9e-25;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 CTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTC 216
|||||
Db 258 CTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTC 317
|||||

QY 217 ACAATCAGCAG 227
|||||
Db 318 ACAATCAGCAG 328
|||||

```

```
RESULT 9
B0946795      918 bp      mRNA      linear      EST 21-AUG-2002
LOCUS          AGENCOURT_8949642 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION     IMAGE:6474845 5', mRNA sequence.
ACCESSION      B0946795
VERSION        B0946795.1 GI:22362273
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus
REFERENCE      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL        National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT       Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14012 row: k column: 06
High quality sequence stop: 714.
Location/Qualifiers
BASE COUNT    246 a 257 c 193 g 212 t 10 others
ORIGIN
Query Match   22.0%; Score 70; DB 14; Length 918;
Best Local Similarity 100.0%; Pred. No. 1.2e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 TGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCA 217
|||||
Db 239 TGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCA 298
|||||

Qy 218 CAATCAGCAG 227
|||||
Db 299 CAATCAGCAG 308
|||||

RESULT 10
B0939046      959 bp      mRNA      linear      EST 21-AUG-2002
LOCUS          AGENCOURT_8946838 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION     IMAGE:6395050 5', mRNA sequence.
ACCESSION      B0939046
VERSION        B0939046.1 GI:22354524
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus
REFERENCE      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL        National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT       Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
```

```
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13889 row: b column: 11
High quality sequence stop: 517.
Location/Qualifiers
FEATURES       source
source         /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone="IMAGE:6395050"
               /clone_lib="NCI_CGAP_Co24"
               /lab_host="DH10B (TI phage-resistant)"
               /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
               Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
               Average insert size 1.6 kb. Constructed by Life
               Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT    267 a 256 c 214 g 222 t
ORIGIN
Query Match   21.4%; Score 68; DB 14; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.3e-23;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 CTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTC 216
|||||
Db 226 CTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTC 285
|||||

Qy 217 ACAATCAG 224
|||||
Db 286 ACAATCAG 293
|||||

RESULT 11
B0968518      798 bp      mRNA      linear      EST 12-JUN-2001
LOCUS          602835104F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4989720 5',
DEFINITION     mRNA sequence.
ACCESSION      B0968518
VERSION        B0968518.1 GI:14356142
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL        National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT       Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11004 row: g column: 01
High quality sequence stop: 797.
Location/Qualifiers
FEATURES       source
source         /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone="IMAGE:4989720"
               /clone_lib="NCI_CGAP_Co24"
               /lab_host="DH10B (TI phage-resistant)"
               /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
```

Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 225 a 222 c 179 g 172 t  
ORIGIN  
Query Match 21.1%; Score 67; DB 13; Length 798;  
Best Local Similarity 100.0%; Pred. No. 4e-23;  
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 CTTCTGGAGTCCCTGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTTCACAA 220  
|||||  
DB 244 CTTCTGGAGTCCCTGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTTCACAA 303  
|||||

QY 221 TCAGCAG 227

DB 304 TCAGCAG 310

RESULT 12  
LOCUS B1150509 755 bp mRNA linear EST 05-JUL-2001  
DEFINITION 602915167f1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:5065720 5',  
mRNA sequence.

ACCESSION B1150509  
VERSION B1150509.1 GI:14610510  
KEYWORDS EST.

SOURCE house mouse.

ORGANISM

REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM1177 row: e column: 17

High quality sequence stop: 738.

FEATURES

source

Location/Qualifiers

1..755

/organism="Mus musculus"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:5065720"

/clone\_lib="NCI\_CGAP\_Lu29"

/tissue\_type="spontaneous tumor, metastatic to mammary.

Stem cell origin.

/lab\_host="DH10B"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 218 a 204 c 168 g 165 t

ORIGIN

Query Match

Best Local Similarity 20.1%; Score 64; DB 13; Length 755;

Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 CTGGAGTCCCTGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTTCACAA 223

DB 247 CTGGAGTCCCTGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTTCACAA 306

QY 224 GCAG 227

||||

DB 307 GCAG 310

RESULT 13

LOCUS B1138873

DEFINITION 601781783f1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4009725 5',

mRNA sequence.

ACCESSION B1138873

VERSION B1138873.1 GI:10977913

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

REFERENCE

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM9246 row: e column: 22

High quality sequence stop: 650.

Location/Qualifiers

1..950

/organism="Mus musculus"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:4009725"

/clone\_lib="NCI\_CGAP\_Lu30"

/tissue\_type="tumor, metastatic to mammary"

/lab\_host="DH10B"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: SalI; transgenic model WNT-1, expression driven by

MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo

dt. Library constructed by Life Technologies.

Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 245 a 277 c 213 g 215 t

ORIGIN

Query Match

Best Local Similarity 19.8%; Score 63; DB 12; Length 950;

Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 CACATCCAACTGGCTTCTGGAGTCCCTGCTTCAGTGGCAGTGGTCTGGGACCTC 206

DB 216 CACATCCAACTGGCTTCTGGAGTCCCTGCTTCAGTGGCAGTGGTCTGGGACCTC 275

QY 207 TTA 209

DB 276 TTA 278

RESULT 14

LOCUS BG149077

DEFINITION 280 bp mRNA linear EST 01-FEB-2001

IMAGE:3183585 5' similar to SW:KVGF\_MOUSE P04940 IG KAPPA CHAIN

V-VI REGION NQ2-17.4.1.; mRNA sequence.

ACCESSION BG149077

VERSION BG149077.1 GI:12652504

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.



REFERENCE 1 (bases 1 to 280)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 MGI:1088581  
 Seq primer: -40RP from Gibco.

FEATURES  
 source Location/Qualifiers  
 1..280

/organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3383585"  
 /clone\_lib="Soares\_mouse\_NMGB\_bcell"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: germinal B-cell; Vector: pT73D-Pac  
 (Pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
 oligo(dT) primer [5',  
 TGGTACCAATCTGAAGTGGAGCGCGCTGGTTTTTTTTTTTTTTTTTTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 Library is normalized; constructed by Bento Soares and  
 M.Fatima Bonaldo."  
 62 a 79 c 70 g 69 t

BASE COUNT  
 ORIGIN

Query Match 18.9%; Score 60; DB 12; Length 280;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 150 ATCCAACTGGCTTCTGGAGTCCCGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 209  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 127 ATCCAACTGGCTTCTGGAGTCCCGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 186

RESULT 15

B0931133  
 LOCUS 918 bp mRNA linear EST 20-AUG-2002  
 DEFINITION AGENCOURT\_8920980 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
 IMAGE:6395584 5', mRNA sequence.  
 B0931133  
 ACCESSION B0931133.1 GI:22346164  
 VERSION B0931133.1  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 918)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13890 row: h column: 17  
 High quality sequence stop: 719.

FEATURES  
 source Location/Qualifiers  
 1..918

/organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6395584"

/clone\_lib="NCI\_CGAP\_Co24"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 230 a 263 c 197 g 216 t 12 others  
 ORIGIN  
 Query Match 18.6%; Score 59; DB 14; Length 918;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-19;  
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 158 TGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTC 216  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 246 TGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTC 304

Search completed: November 27, 2002, 07:13:38  
 Job time : 1333.94 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:55:59 ; Search time 35.6419 Seconds  
(without alignments)  
2736.194 Million cell updates/sec

Title: US-09-893-615-88  
Perfect score: 318  
Sequence: 1 CAATGTGTTCTCTCCAGTC.....GGACCATGCTGGAATAAGA 318

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0  
Searched: 441362 seqs, 15338381 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2.6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	54.1	318	3 US-08-783-853A-104	Sequence 104, App
2	172	54.1	318	4 US-09-344-050-104	Sequence 104, App
3	172	54.1	321	3 US-08-783-853A-6	Sequence 6, Appli
4	172	54.1	321	4 US-09-344-050-6	Sequence 6, Appli
5	172	54.1	335	3 US-08-783-853A-102	Sequence 102, App
6	172	54.1	335	4 US-09-344-050-102	Sequence 102, App
7	120	37.7	387	2 US-08-449-287-1	Sequence 1, Appli
8	120	37.7	705	4 US-09-423-439-17	Sequence 17, Appl
9	120	37.7	732	2 US-08-860-882A-26	Sequence 26, Appl
10	120	37.7	732	4 US-09-423-439-57	Sequence 57, Appl
11	120	37.7	732	4 US-09-011-769A-22	Sequence 22, Appl
12	120	37.7	3217	4 US-09-423-439-52	Sequence 52, Appl
13	90	28.3	222	2 US-08-308-494A-12	Sequence 12, Appl
14	90	28.3	232	4 US-09-280-028-3	Sequence 3, Appli
15	90	28.3	315	1 US-08-459-310-3	Sequence 3, Appli
16	71	22.3	300	2 US-08-308-494A-20	Sequence 20, Appl
17	71	22.3	309	1 US-08-467-393-3	Sequence 3, Appli
18	71	22.3	331	3 US-08-836-561-32	Sequence 32, Appl
19	71	22.3	384	2 US-08-656-586-1	Sequence 1, Appli
20	71	22.3	384	4 US-08-619-491-1	Sequence 1, Appli
21	71	22.3	384	5 PCT-US95-07302-1	Sequence 1, Appli
22	70	22.0	360	1 US-08-447-422-13	Sequence 13, Appl
23	70	22.0	648	6 5455030-4	Patent No. 5455030
24	70	22.0	788	6 5455030-6	Patent No. 5455030
25	70	22.0	786	4 US-08-635-928-31	Sequence 31, Appl
26	70	22.0	1848	1 US-08-447-422-15	Sequence 15, Appl
27	68	21.4	711	6 5455030-8	Patent No. 5455030

28	68	21.4	741	6	5455030-10	Patent No. 5455030
29	67	21.1	384	1	US-08-149-099C-6	Sequence 6, Appli
30	67	21.1	384	2	US-08-478-967A-6	Sequence 6, Appli
31	67	21.1	384	4	US-08-475-815B-6	Sequence 6, Appli
32	67	21.1	672	2	US-08-190-199A-62	Sequence 62, Appl
33	67	21.1	711	2	US-08-190-199A-64	Sequence 64, Appl
34	67	21.1	738	2	US-08-956-047-24	Sequence 24, Appl
35	67	21.1	784	2	US-08-956-047-32	Sequence 32, Appl
36	67	21.1	793	2	US-08-956-047-29	Sequence 29, Appl
37	67	21.1	1256	4	US-09-553-498-7	Sequence 7, Appli
38	67	21.1	1256	4	US-09-618-869-7	Sequence 7, Appli
39	67	21.1	9209	1	US-08-149-099C-3	Sequence 3, Appli
40	67	21.1	9209	2	US-08-478-967A-3	Sequence 3, Appli
41	67	21.1	9209	4	US-08-475-815B-3	Sequence 3, Appli
42	67	21.1	18986	2	US-08-819-866-2	Sequence 2, Appli
43	67	21.1	18986	2	US-09-023-715-2	Sequence 2, Appli
44	67	21.1	18986	4	US-09-343-485A-2	Sequence 2, Appli
45	63	19.8	384	3	US-08-444-644-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-08-783-853A-104  
; Sequence 104, Application US/08783853A  
; Patent No. 6005091  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; APPLICANT: Church, William  
; APPLICANT: Gross, Mitchell  
; APPLICANT: Feuerstein, Giora  
; APPLICANT: Nichols, Andrew  
; APPLICANT: Padlan, Eduardo  
; APPLICANT: Patel, Arunbhai  
; APPLICANT: Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/783,853A  
; FILING DATE: 16-JAN-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/029,119  
; FILING DATE: 24-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 104:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 318 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA

```
;
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...318
; OTHER INFORMATION:
US-08-783-853A-104

Query Match      54.1%; Score 172; DB 3; Length 318;
Best Local Similarity 99.6%; Pred. No. 2.le-80;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
Qy 112 AAGCCAGGATCCTCCCCAAACCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171
Db 112 AAGCCAGGATCCTCCCCAAACCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171
Qy 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
Qy 232 GAGGCTGAAGATGCTGCCACCTTATTACTGCCAGCAGTGGAGTA 274
Db 232 GAGGCTGAAGATGCTGCCACCTTATTACTGCCAGCAGTGGAGTA 274

RESULT 2
US-09-344-050-104
; Sequence 104, Application US/09344050
; Patent No. 6391299
; GENERAL INFORMATION:
; APPLICANT: Blackburn, Michael
; APPLICANT: Church, William
; APPLICANT: Gross, Mitchell
; APPLICANT: Feuerstein, Giora
; APPLICANT: Nichols, Andrew
; APPLICANT: Padlan, Eduardo
; APPLICANT: Patel, Arunbhai
; APPLICANT: Sylvester, Daniel
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,050
; FILING DATE: 24-JUN-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/783,853
; FILING DATE: 16-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: P50438
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX:
```

```
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...318
; OTHER INFORMATION:
US-09-344-050-104

Query Match      54.1%; Score 172; DB 4; Length 318;
Best Local Similarity 99.6%; Pred. No. 2.le-80;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
Db 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111
Qy 112 AAGCCAGGATCCTCCCCAAACCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171
Db 112 AAGCCAGGATCCTCCCCAAACCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171
Qy 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231
Qy 232 GAGGCTGAAGATGCTGCCACCTTATTACTGCCAGCAGTGGAGTA 274
Db 232 GAGGCTGAAGATGCTGCCACCTTATTACTGCCAGCAGTGGAGTA 274

RESULT 3
US-08-783-853A-6
; Sequence 6, Application US/08783853A
; Patent No. 6005091
; GENERAL INFORMATION:
; APPLICANT: Blackburn, Michael
; APPLICANT: Church, William
; APPLICANT: Gross, Mitchell
; APPLICANT: Feuerstein, Giora
; APPLICANT: Nichols, Andrew
; APPLICANT: Padlan, Eduardo
; APPLICANT: Patel, Arunbhai
; APPLICANT: Sylvester, Daniel
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/783,853A
; FILING DATE: 16-JAN-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/029,119
```

FILING DATE: 24-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-783-853A-6

Query Match 54.1%; Score 172; DB 3; Length 321;  
Best Local Similarity 99.6%; Pred. No. 2.le-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACCTTGCAGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
DB 52 AAGGTCACAATGACCTTGCAGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCCTCCCAAAACCCCTGGATTTCTGCCACATCCAACTGGCTTCTTGAGTC 171  
|||||  
DB 112 AAGCCAGGATCCTCCCAAAACCCCTGGATTTATGCCACATCCAACTGGCTTCTTGAGTC 171  
|||||  
QY 172 CCGTCTCGCTTCAGTGGCAGTGGGCTCGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
|||||  
DB 172 CCGTCTCGCTTCAGTGGCAGTGGGCTCGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

RESULT 4  
US-09-344-050-6  
Sequence 6, Application US/09344050  
Patent No. 6391299

GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Giora  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,050

FILING DATE: 24-JUN-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-09-344-050-6

Query Match 54.1%; Score 172; DB 4; Length 321;  
Best Local Similarity 99.6%; Pred. No. 2.le-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACCTTGCAGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
DB 52 AAGGTCACAATGACCTTGCAGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCCTCCCAAAACCCCTGGATTTCTGCCACATCCAACTGGCTTCTTGAGTC 171  
|||||  
DB 112 AAGCCAGGATCCTCCCAAAACCCCTGGATTTATGCCACATCCAACTGGCTTCTTGAGTC 171  
|||||  
QY 172 CCGTCTCGCTTCAGTGGCAGTGGGCTCGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
|||||  
DB 172 CCGTCTCGCTTCAGTGGCAGTGGGCTCGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

RESULT 5  
US-08-783-853A-102  
Sequence 102, Application US/08783853A  
Patent No. 6005091

GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Giora  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/783,853A  
FILING DATE: 16-JAN-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/029,119  
FILING DATE: 24-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:

## INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:  
LENGTH: 335 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:

NAME/KEY: Coding Sequence  
LOCATION: 1...335  
OTHER INFORMATION:  
US-08-783-853A-102

Query Match 54.1%; Score 172; DB 3; Length 335;

Best Local Similarity 99.6%; Pred. No. 2.1e-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
DB 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCTCCCAACCCCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171  
|||||  
DB 112 AAGCCAGGATCTCCCAACCCCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171  
|||||  
QY 172 CCGTCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
DB 172 CCGTCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

## RESULT 6

US-09-344-050-102  
Sequence 102, Application US/09344050  
Patent No. 6391299

## GENERAL INFORMATION:

APPLICANT: Blackburn, Michael  
APPLICANT: Church, William  
APPLICANT: Gross, Mitchell  
APPLICANT: Feuerstein, Giora  
APPLICANT: Nichols, Andrew  
APPLICANT: Padlan, Eduardo  
APPLICANT: Patel, Arunbhai  
APPLICANT: Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
TITLE OF INVENTION: OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,050  
FILING DATE: 24-JUN-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumelster, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:

## INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:  
LENGTH: 335 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...335  
OTHER INFORMATION:  
US-09-344-050-102

Query Match 54.1%; Score 172; DB 4; Length 335;

Best Local Similarity 99.6%; Pred. No. 2.1e-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
DB 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAATATACATGCACCTGGTACCAGCAG 111  
|||||  
QY 112 AAGCCAGGATCTCCCAACCCCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171  
|||||  
DB 112 AAGCCAGGATCTCCCAACCCCTGGATTTCGCCACATCCAACTGGCTTCTGGAGTC 171  
|||||  
QY 172 CCGTCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
DB 172 CCGTCTCGCTTCAGTGGCAGTGGGTCTGGACCTCTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

## RESULT 7

US-08-449-287-1  
Sequence 1, Application US/08449287  
Patent No. 5877293

## GENERAL INFORMATION:

APPLICANT: ADAIR, John Robert  
APPLICANT: BODMER, Mark William  
APPLICANT: MOUNTAIN, Andrew  
APPLICANT: OWENS, Raymond John  
TITLE OF INVENTION: CDR Grafted Anti-CEA Antibodies and

TITLE OF INVENTION: Their Production  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,287  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/154,389  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT GB91/01108  
FILING DATE: 05-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9014932.9  
FILING DATE: 05-JUL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT GB90/02017  
FILING DATE: 21-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: SAXE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 40283/110 CARA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
CLONE: A5B7 light chain V1 domain  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..387  
US-08-449-287-1

Query Match 37.7%; Score 120; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 2.6e-53;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 144 TGCCACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGAC 203  
|||||  
Db 210 TGCCACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGAC 269  
QY 204 CTCCTACTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 263  
|||||  
Db 270 CTCCTACTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 329

RESULT 8  
US-09-423-439-17  
Sequence 17, Application US/09423439  
Patent No. 6339070  
GENERAL INFORMATION:  
APPLICANT: EMERY, Stephen Charles  
TITLE OF INVENTION: CHEMICAL COMPOUNDS  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pillsbury Winthrop, L.L.P.  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/423,439  
FILING DATE: 09-No. 6339070-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB98/01294  
FILING DATE: 05-MAY-1998  
APPLICATION NUMBER: GB 9709421.3  
FILING DATE: 10-MAY-1997  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 705 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-423-439-17

Query Match 37.7%; Score 120; DB 4; Length 705;  
Best Local Similarity 100.0%; Pred. No. 2.7e-53;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 144 TGCCACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGAC 203  
|||||  
Db 210 TGCCACATCCAACTGGCTTCTGGAGTCCCTCGCTTCAGTGGCAGTGGGTCTGGGAC 269  
QY 204 CTCCTACTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 263  
|||||  
Db 270 CTCCTACTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 329

RESULT 9  
US-08-860-882A-26  
Sequence 26, Application US/08860882A  
Patent No. 5985281  
GENERAL INFORMATION:  
APPLICANT: TAYLORSON, CHRISTOPHER JOHN  
APPLICANT: EGGELE, HENDRIKUS JOHANNES  
APPLICANT: TARRAGONA-FIOL, ANTONIO  
APPLICANT: RABIN, BRIAN ROBERT  
APPLICANT: BOYLE, FRANCIS THOMAS  
APPLICANT: HENNAM, JOHN FREDERICK  
APPLICANT: BLAKELY, DAVID CHARLES  
APPLICANT: MARSHAM, PETER ROBERT  
APPLICANT: HEATON, DAVID WILLIAM  
APPLICANT: DAVIES, DAVID HUW  
TITLE OF INVENTION: CHEMICAL COMPOUNDS  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY, MADISON & SUTRO  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/860,882A  
FILING DATE: JUNE 23, 1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: DONALD J. BIRD  
REGISTRATION NUMBER: 25,323  
REFERENCE/DOCKET NUMBER: 9901/238653  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3027  
TELEFAX: (202) 822-0944  
TELEX: 6174627 CUSH  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-860-882A-26

Query Match 37.7%; Score 120; DB 2; Length 732;  
Best Local Similarity 100.0%; Pred. No. 2.7e-53;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 203  
Db 225 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 284  
Qy 204 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 263  
Db 285 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 344

RESULT 10  
US-09-423-439-57  
Sequence 57, Application US/09423439  
Patent No. 6339070  
GENERAL INFORMATION:  
APPLICANT: EMERY, Stephen Charles  
BLAKEY, David Charles  
TITLE OF INVENTION: CHEMICAL COMPOUNDS  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Winthrop, L.L.P.  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/423,439  
FILING DATE: 09-NOV-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB98/01294  
FILING DATE: 05-MAY-1998  
APPLICATION NUMBER: GB 9709421.3  
FILING DATE: 10-MAY-1997  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16..720  
SEQUENCE DESCRIPTION: SEQ ID NO: 57:

Qy 144 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 203  
Db 225 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 284  
Qy 204 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 263  
Db 285 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 344

US-09-423-439-57  
Query Match 37.7%; Score 120; DB 4; Length 732;  
Best Local Similarity 100.0%; Pred. No. 2.7e-53;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 203  
Db 225 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 284  
Qy 204 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 263  
Db 285 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 344

RESULT 11  
US-09-011-769A-22  
Sequence 22, Application US/09011769A  
Patent No. 6436691  
GENERAL INFORMATION:  
APPLICANT: SLATER, Anthony M.  
BLAKEY, David C.  
DAVIES, David H.  
HENNAM, John F.  
HENNEQUIN, Laurent F.A.  
MARSHAM, Peter R.  
DOWELL, Robert I.  
TITLE OF INVENTION: Chemical Compounds  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Madison & Sutro, LLP  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 Mb disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/011,769A  
FILING DATE: 13-Feb-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB96/01975  
FILING DATE: 13-AUG-1996  
APPLICATION NUMBER: GB 9612295.7  
FILING DATE: 12-JUN-1996  
APPLICATION NUMBER: GB 9611019.2  
FILING DATE: 25-MAY-1996  
APPLICATION NUMBER: GB 9516810.0  
FILING DATE: 16-AUG-1995  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16..720  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-011-769A-22

Qy 144 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 203  
Db 225 TGCACATCCAACCTGCTTCTGAGTCCCTGCTGCTTTCAGTGGCAGTGGTCTGGGAC 284  
Qy 204 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 263  
Db 285 CTCCTACTCTCTCAATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCA 344



Db 225 TGCCACATCCAACTGGCTTCTGGAGTCCCTGCTGCTTTCAGTGGCAGTGGGTCTGGGAC 284  
QY 204 CTCTTACTCTCTCACATCAGCAGAGTGGAGCTGAAGATGCTGCCACTTATTACTGCCA 263  
Db 285 CTCTTACTCTCTCACATCAGCAGAGTGGAGCTGAAGATGCTGCCACTTATTACTGCCA 344

## RESULT 12

US-09-423-439-52

; Sequence 52, Application US/09423439

; Patent No. 6339070

; GENERAL INFORMATION:

; APPLICANT: EMERY, Stephen Charles

; BLAKEY, David Charles

; TITLE OF INVENTION: CHEMICAL COMPOUNDS

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pillsbury Winthrop, L.L.P.

; STREET: 1100 New York Ave., N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: MS Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/423,439

; FILING DATE: 09-NOV-1999

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB98/01294

; FILING DATE: 05-MAY-1998

; APPLICATION NUMBER: GB 9709421.3

; FILING DATE: 10-MAY-1997

; INFORMATION FOR SEQ ID NO: 52:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3217 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; SEQUENCE DESCRIPTION: SEQ ID NO: 52:

US-09-423-439-52

## Query Match

Best Local Similarity 37.7%; Score 120; DB 4; Length 3217;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 TGCCACATCCAACTGGCTTCTGGAGTCCCTGCTGCTTTCAGTGGCAGTGGGTCTGGGAC 203

Db 225 TGCCACATCCAACTGGCTTCTGGAGTCCCTGCTGCTTTCAGTGGCAGTGGGTCTGGGAC 284

QY 204 CTCTTACTCTCTCACATCAGCAGAGTGGAGCTGAAGATGCTGCCACTTATTACTGCCA 263

Db 285 CTCTTACTCTCTCACATCAGCAGAGTGGAGCTGAAGATGCTGCCACTTATTACTGCCA 344

## RESULT 13

US-08-308-494A-12

; Sequence 12, Application US/08308494A

; Patent No. 5959083

; GENERAL INFORMATION:

; APPLICANT: Bosslet, Klaus

; APPLICANT: Seeman, Gerhard

; TITLE OF INVENTION: Tetraivalent Bispecific Receptors, The

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &amp;

; ADDRESS: Dunner

; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/308,494A  
; FILING DATE: 21-SEP-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P4118120.4  
; FILING DATE: 03-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kulik, David J.  
; REGISTRATION NUMBER: 36,576  
; REFERENCE/DOCKET NUMBER: 05552-1186-02000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 292 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-308-494A-12

## Query Match

Best Local Similarity 28.3%; Score 90; DB 2; Length 292;

Matches 190; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 90 TTACATGTCACCTGGTGTACACAGAGCCAGGATCCTCCCCAAACCTGGATTCTGCCAC 149

Db 67 TTACATGTCACCTGGTGTACACAGAGCCAGGATCCTCCCCAAACCTGGATTCTGCCAC 126

QY 150 ATCCAACTGGCTTCTGGAGTCCCTGCTGCTTTCAGTGGCAGTGGGTCTGGGACCTCTTA 209

Db 127 ATCCAACTGGCTTCTGGAGTCCCTGCTGCTTTCAGTGGCAGTGGGTCTGGGACCTCTTA 186

QY 210 CTCTCTCACATCAGCAGAGTGGAGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269

Db 187 CTCTCTCACATCAGCAGAGTGGAGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 246

QY 270 GAGTAGTAACCC 281

Db 247 GAGTAGTAACCC 258

## RESULT 14

US-09-280-028-3

; Sequence 3, Application US/09280028

; Patent No. 6241961

; GENERAL INFORMATION:

; APPLICANT: BENES, IVAN FRIEDRICH

; APPLICANT: BOSSLET, KLAUS

; TITLE OF INVENTION: RADIOIMMUNO CONJUGATES FOR USE IN HUMAN THERAPY AND

; FILE REFERENCE: BENES

; CURRENT APPLICATION NUMBER: US/09/280,028

; CURRENT FILING DATE: 1999-03-26

; EARLIER APPLICATION NUMBER: 199 11 329.7

; EARLIER FILING DATE: 1999-03-15

; EARLIER APPLICATION NUMBER: 198 13 687.0

```
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: cDNA sequence of codons for amino acids 9-105 of
; OTHER INFORMATION: VL exon for k-light chain gene (VK) which codes
; OTHER INFORMATION: for the MAb light chain and the amino acid
; OTHER INFORMATION: sequence associated therewith
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(292)
US-09-280-028-3

Query Match      28.3%; Score 90; DB 4; Length 292;
Best Local Similarity 99.0%; Pred. No. 1.1e-37;
Matches 190; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 90 TTACATGACACTGGTACCAGCAGAGAGCCAGGATCCTCCCCAAACCCCTGGATTCTGCCAC 149
Db 67 TTACATGACACTGGTACCAGCAGAGAGCCAGGATCCTCCCCAAACCCCTGGATTCTGCCAC 126

Qy 150 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 209
Db 127 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 186

Qy 210 CTCCTCTCACATATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
Db 187 CTCCTCTCACATATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 246

Qy 270 GAGTAGTAACCC 281
Db 247 GAGTAGTAACCC 258

RESULT 15
US-08-459-310-3
; Sequence 3, Application US/08459310
; Patent No. 5645817
; GENERAL INFORMATION:
; APPLICANT: Seemann, Gerhard
; APPLICANT: Bosslet, Klaus
; TITLE OF INVENTION: Granulocyte-Binding Antibody Constructs,
; TITLE OF INVENTION: Their Preparation and Use
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,310
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,963
; FILING DATE: 03-AUG-1993
; APPLICATION NUMBER: DE P 422 58 53.7
; FILING DATE: 05-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wadler, Linda A.
; REGISTRATION NUMBER: 33,218
```

```
; REFERENCE/DOCKET NUMBER: 02481.1317-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-459-310-3

Query Match      28.3%; Score 90; DB 1; Length 315;
Best Local Similarity 99.0%; Pred. No. 1.1e-37;
Matches 190; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 90 TTACATGACACTGGTACCAGCAGAGCCAGGATCCTCCCCAAACCCCTGGATTCTGCCAC 149
Db 90 TTACATGACACTGGTACCAGCAGAGCCAGGATCCTCCCCAAACCCCTGGATTCTGCCAC 149

Qy 150 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 209
Db 150 ATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTA 209

Qy 210 CTCCTCTCACATATCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269
Db 210 CTCCTCTCACATATCAGCAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTG 269

Qy 270 GAGTAGTAACCC 281
Db 270 GAGTAGTAACCC 281
```

Search completed: November 27, 2002, 05:39:27  
Job time : 39.6419 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 05:34:00 ; Search time 35.6419 Seconds  
(without alignments)  
3436.030 Million cell updates/sec

Title: US-09-893-615-88

Perfect score: 318

Sequence: 1 CAAATGTTCTCTCCAGTC.....GGACCATGCTGGAATAAGA 318

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 341543 seqs, 192557720 residues

Word size : 0

Total number of hits satisfying chosen parameters: 693086

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	318	100.0	318	10	US-09-893-615-88
2	172	54.1	318	10	US-09-965-099-104
3	172	54.1	318	12	US-10-051-852-104
4	172	54.1	321	10	US-09-965-099-6
5	172	54.1	321	12	US-10-051-852-6
6	172	54.1	335	10	US-09-965-099-102
7	172	54.1	335	12	US-10-051-852-102
8	55	17.3	390	10	US-09-753-436-79
9	53	16.7	165	10	US-09-965-099-63
10	53	16.7	165	12	US-10-051-852-63
11	53	16.7	280	10	US-09-965-099-65
12	53	16.7	280	12	US-10-051-852-65
13	53	16.7	321	10	US-09-965-099-73
14	53	16.7	321	12	US-10-051-852-73
15	53	16.7	357	10	US-09-808-037-27
16	53	16.7	717	8	US-08-940-344-3
17	53	16.7	717	10	US-09-808-037-5
18	51	16.0	5691	10	US-09-897-006-11
19	48	15.1	318	10	US-09-976-787-27

20	48	15.1	318	10	US-09-865-198-26
21	48	15.1	322	9	US-09-982-107-11
22	48	15.1	324	10	US-09-976-787-16
23	48	15.1	324	10	US-09-865-198-16
24	48	15.1	357	10	US-09-910-059-8
25	48	15.1	705	10	US-09-910-059-16
26	48	15.1	714	10	US-09-976-787-31
27	48	15.1	720	10	US-09-976-787-30
28	48	15.1	2059	9	US-09-807-721-1
29	44	13.8	330	10	US-09-965-099-94
30	44	13.8	330	12	US-10-051-852-94
31	44	13.8	412	10	US-09-965-099-98
32	44	13.8	412	12	US-10-051-852-98
33	41	12.9	165	10	US-09-965-099-32
34	41	12.9	165	12	US-10-051-852-32
35	41	12.9	280	10	US-09-965-099-34
36	41	12.9	280	12	US-10-051-852-34
37	41	12.9	321	10	US-09-965-099-43
38	41	12.9	321	10	US-09-965-099-56
39	41	12.9	321	10	US-09-965-099-61
40	41	12.9	321	12	US-10-051-852-43
41	41	12.9	321	12	US-10-051-852-56
42	41	12.9	321	12	US-10-051-852-61
43	41	12.9	401	10	US-09-965-099-77
44	41	12.9	401	12	US-10-051-852-77
45	36	11.3	108	10	US-09-965-099-91

#### ALIGNMENTS

#### RESULT 1

US-09-893-615-88

; Sequence 88 Application US/09893615

; Patent No. US20020082395A1

; GENERAL INFORMATION:

APPLICANT: Fischer, Gerald W.

Schuman, Richard F.

Wong, Hing

Stinson, Jeffrey L.

TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND

CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR

POSITIVE BACTERIA

NUMBER OF SEQUENCES: 89

CORRESPONDENCE ADDRESS:

ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/893,615

FILING DATE: 29-Jun-2001

ATTORNEY/AGENT INFORMATION:

NAME: Elnaudi, Carol P.

REGISTRATION NUMBER: 32,220

REFERENCE/DOCKET NUMBER: 04995.0041-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 88:

SEQUENCE CHARACTERISTICS:

LENGTH: 318 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

```

; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..318
; SEQUENCE DESCRIPTION: SEQ ID NO: 88:
US-09-893-615-88

Query Match          100.0%; Score 318; DB 10; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.7e-155; Indels 0; Gaps 0;
Matches 318; Conservative 0; Mismatches 0;

QY 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAAAAGGTTCACA 60
   |||||||
Db 1 CAAATTGTTCTCTCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAAAAGGTTCACA 60
   |||||||

QY 61 ATGACTTGAGGCCAGCCAGCTCAAGTGTAAATTACATGCATCTGGTACCAGCAGAGCCAGGA 120
   |||||||
Db 61 ATGACTTGAGGCCAGCCAGCTCAAGTGTAAATTACATGCATCTGGTACCAGCAGAGCCAGGA 120
   |||||||

QY 121 TCTCCCCCAAAACCCCTGGGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180
   |||||||
Db 121 TCTCCCCCAAAACCCCTGGGATTTCTGCCACATCCAACTGGCTTCTGGAGTCCCTGCTCGC 180
   |||||||

QY 181 TTCAGTGGCAGTGGGTCCTGGGACCTCTTACTCTCTCACAATCAGCAGAGCTGGAGGGCTGAA 240
   |||||||
Db 181 TTCAGTGGCAGTGGGTCCTGGGACCTCTTACTCTCTCACAATCAGCAGAGCTGGAGGGCTGAA 240
   |||||||

QY 241 GATGCTGCCACTATTACTGCCAGCAGTGGAGTAGTAACCCACCCACCTTCGAGGGGGG 300
   |||||||
Db 241 GATGCTGCCACTATTACTGCCAGCAGTGGAGTAGTAACCCACCCACCTTCGAGGGGGG 300
   |||||||

QY 301 ACCATGCTGGAATAAGA 318
   |||||||
Db 301 ACCATGCTGGAATAAGA 318
   |||||||

RESULT 2
US-09-965-099-104
; Sequence 104, Application US/09965099
; Patent No. US20020136725A1
; GENERAL INFORMATION:
; APPLICANT: Blackburn, Michael
; Feuerstein, Giora
; Patel, Arunbhai
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN
; TREATMENT OF THROMBOSIS
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/965,099
; FILING DATE: 26-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/346,487
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: P50438-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: <Unknown>

```

```

; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1..318
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-965-099-104

Query Match          54.1%; Score 172; DB 10; Length 318;
Best Local Similarity 99.6%; Pred. No. 8.1e-80;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTTCAGAGGCCAGCTCAAGTGTAAATTACATGCATCTGGTACCAGCAG 111
   |||||||
Db 52 AAGGTCACAATGACTTTCAGAGGCCAGCTCAAGTGTAAATTACATGCATCTGGTACCAGCAG 111
   |||||||

QY 112 AAGCCAGGAGTCTCCCCCAACCCCTGGATTTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171
   |||||||
Db 112 AAGCCAGGAGTCTCCCCCAACCCCTGGATTTCTGCCACATCCCAACCTGGCTTCTGGAGTC 171
   |||||||

QY 172 CTTGCTCGCTTCAGTGGCAGTGGGTCTGGAGCTCTTACTCTCTCACAATCAGCAGAGTG 231
   |||||||
Db 172 CTTGCTCGCTTCAGTGGCAGTGGGTCTGGAGCTCTTACTCTCTCACAATCAGCAGAGTG 231
   |||||||

QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274
   |||||||
Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274
   |||||||

RESULT 3
US-10-051-852-104
; Sequence 104, Application US/10051852
; Patent No. US2002014641A1
; GENERAL INFORMATION:
; APPLICANT: Blackburn, Michael
; Church, William
; Gross, Mitchell
; Feuerstein, Giora
; Nicholls, Andrew
; Padlan, Eduardo
; Patel, Arunbhai
; Sylvester, Daniel
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT
; OF THROMBOSIS
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/051,852
; FILING DATE: 17-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...318  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 104:  
US-10-051-852-104

Query Match 54.1%; Score 172; DB 12; Length 318;  
Best Local Similarity 99.6%; Pred. No. 8.1e-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 52 AAGGTCAACATGACTTGCGAGGCCAGCTCAAGTGTAAATTACATGCATCGGTACCAGCAG 111  
Db 52 AAGGTCAACATGACTTGCGAGGCCAGCTCAAGTGTAAATTACATGCATCGGTACCAGCAG 111

Qy 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171  
Db 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171

Qy 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCACAATCAGCAGAGTG 231  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCACAATCAGCAGAGTG 231

Qy 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274

RESULT 4  
US-09-965-099-6  
Sequence 6, Application US/09965099  
Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Giora  
Patel, Arunbhai  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-965-099-6

Query Match 54.1%; Score 172; DB 10; Length 321;  
Best Local Similarity 99.6%; Pred. No. 8.1e-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 52 AAGGTCAACATGACTTGCGAGGCCAGCTCAAGTGTAAATTACATGCATCGGTACCAGCAG 111  
Db 52 AAGGTCAACATGACTTGCGAGGCCAGCTCAAGTGTAAATTACATGCATCGGTACCAGCAG 111

Qy 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171  
Db 112 AAGCCAGGATCTCCCCAAACCCCTGGATTCTGCCACATCCAACTGGCTTCTGGAGTC 171

Qy 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCACAATCAGCAGAGTG 231  
Db 172 CTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTCTACTCTCTCACAATCAGCAGAGTG 231

Qy 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
Db 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274

RESULT 5  
US-10-051-852-6  
Sequence 6, Application US/10051852  
Patent No. US20020146411A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Giora  
Nichols, Andrew  
Padlan, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA

ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-051-852-6

Query Match 54.1%; Score 172; DB 12; Length 321;  
Best Local Similarity 99.6%; Pred. No. 8.1e-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111  
|||||  
DB 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111  
|||||

QY 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTTGCACACATCCAACTGCTCTGGAGTC 171  
|||||  
DB 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTTGCACACATCCAACTGCTCTGGAGTC 171  
|||||

QY 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
DB 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCAATCAGCAGAGTG 231  
|||||

QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

## RESULT 6

US-09-965-099-102  
; Sequence 102, Application US/09965099  
; Patent No. US20020136725A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Feuerstein, Giora  
; Patel, Arunbhai  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
; TREATMENT OF THROMBOSIS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia

STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 335 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...335  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 102:  
US-09-965-099-102

Query Match 54.1%; Score 172; DB 10; Length 335;  
Best Local Similarity 99.6%; Pred. No. 8.1e-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111  
|||||  
DB 52 AAGGTCACAATGACTTGCAGGGCCAGCTCAAGTGTAAATTACATGCACTGGTACCAGCAG 111  
|||||

QY 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTTGCACACATCCAACTGCTCTGGAGTC 171  
|||||  
DB 112 AAGCCAGGATCTCCGCCAAACCTGGATTCTTGCACACATCCAACTGCTCTGGAGTC 171  
|||||

QY 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCAATCAGCAGAGTG 231  
|||||  
DB 172 CCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTTACTCTCTCAATCAGCAGAGTG 231  
|||||

QY 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||  
DB 232 GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTA 274  
|||||

## RESULT 7

US-10-051-852-102  
; Sequence 102, Application US/10051852  
; Patent No. US2002014641A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Church, William  
; Gross, Mitchell  
; Feuerstein, Giora  
; Nichols, Andrew  
; Padlan, Eduardo

Patel, Arunbhai  
Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/051.852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 335 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...335  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 102:  
US-10-051-852-102

Query Match 54.1%; Score 172; DB 12; Length 335;  
Best Local Similarity 99.6%; Pred. No. 8.le-80;  
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 52 AAGGTCACAATGACTTGCAGGCCAGCTCAAGTGTAAATTACATGCTGTACCGAG 111  
Db 52 AAGGTCACAATGACTTGCAGGCCAGCTCAAGTGTAAATTACATGCTGTACCGAG 111  
Qy 112 AAGCCAGGATCTCCCCCAAAACCTTGGATTTTGCACATCAACCTGGCTTCTGGAGTC 171  
Db 112 AAGCCAGGATCTCCCCCAAAACCTTGGATTTTGCACATCAACCTGGCTTCTGGAGTC 171  
Qy 172 CCGTCTCGCTTCAAGTGGCAGTGGGCTCGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
Db 172 CCGTCTCGCTTCAAGTGGCAGTGGGCTCGGACCTCTTACTCTCTCACAATCAGCAGAGTG 231  
Qy 232 GAGGCTGAAGATGCTGCCACCTATTACTGCCAGCAGTGGAGTA 274  
Db 232 GAGGCTGAAGATGCTGCCACCTATTACTGCCAGCAGTGGAGTA 274

RESULT 8  
US-09-753-436-79  
Sequence 79, Application US/09753436  
Patent No. US20010029293A1  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemay  
TITLE OF INVENTION: ICAM-Related Materials and Methods  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/753.436  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/382,289  
FILING DATE:  
APPLICATION NUMBER: US 08/487,113  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,754  
FILING DATE: 05-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Joseph A., Jr.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 33282  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 390 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-753-436-79

Query Match 17.3%; Score 55; DB 10; Length 390;  
Best Local Similarity 100.0%; Pred. No. 3.6e-19;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 CACATCCAACCTGGCTTCTGGAGTCCCTGCTTCAGTGGCAGTGGTCTGGG 201  
Db 219 CACATCCAACCTGGCTTCTGGAGTCCCTGCTTCAGTGGCAGTGGTCTGGG 273

## RESULT 9

US-09-965-099-63  
; Sequence 63, Application US/09965099  
; Patent No. US20020136725A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Feuerstein, Giora  
; Patel, Arunbhai  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
; TREATMENT OF THROMBOSIS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/965,099  
; FILING DATE: 26-Sep-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/346,487  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: <Unknown>  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 165 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:  
US-09-965-099-63

Query Match 16.7%; Score 53; DB 10; Length 165;  
Best Local Similarity 100.0%; Pred. NO. 3.8e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTACCGAGGCGCAGG 119  
Db 62 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTGTTACCGAGGCGCAGG 114

## RESULT 10

US-10-051-852-63  
; Sequence 63, Application US/10051852  
; Patent No. US2002014641A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Church, William  
; Gross, Mitchell  
; Feuerstein, Giora  
; Nichols, Andrew  
; Padlan, Eduardo

Patel, Arunbhai  
Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; OF THROMBOSIS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/051,852  
; FILING DATE: 17-Jan-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/344,050  
; FILING DATE: 25-JUN-1999  
; APPLICATION NUMBER: 08/783,853  
; FILING DATE: 16-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: <Unknown>  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 165 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:  
US-10-051-852-63

Query Match 16.7%; Score 53; DB 12; Length 165;  
Best Local Similarity 100.0%; Pred. NO. 3.8e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTACCGAGGCGCAGG 119  
Db 62 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTGTTACCGAGGCGCAGG 114

## RESULT 11

US-09-965-099-65  
; Sequence 65, Application US/09965099  
; Patent No. US20020136725A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Feuerstein, Giora  
; Patel, Arunbhai  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
; TREATMENT OF THROMBOSIS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA



;  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/965,099  
; FILING DATE: 26-Sep-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/346,487  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: <Unknown>  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 2...280  
; OTHER INFORMATION:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-09-965-099-65

Query Match 16.7%; Score 53; DB 10; Length 280;  
Best Local Similarity 100.0%; Pred. No. 3.8e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTACACAGCAGG 119  
Db 62 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTACACAGCAGG 114

RESULT 12  
US-10-051-852-65  
; Sequence 65, Application US/10051852  
; Patent No. US2002014611A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Church, William  
; Gross, Mitchell  
; Feuerstein, Giora  
; Nichols, Andrew  
; Padlan, Eduardo  
; Patel, Arunbhai  
; Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; OF THROMBOSIS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:

;  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/051,852  
; FILING DATE: 17-Jan-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/344,050  
; FILING DATE: 25-JUN-1999  
; APPLICATION NUMBER: 08/783,853  
; FILING DATE: 16-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: <Unknown>  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 2...280  
; OTHER INFORMATION:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-10-051-852-65

Query Match 16.7%; Score 53; DB 12; Length 280;  
Best Local Similarity 100.0%; Pred. No. 3.8e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTACACAGCAGG 119  
Db 62 TGCAGGCCAGCTCAAGTGTAAATTACATGCTGCTACACAGCAGG 114

RESULT 13  
US-09-965-099-73  
; Sequence 73, Application US/09965099  
; Patent No. US20020136725A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Feuerstein, Giora  
; Patel, Arunbhai  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
; TREATMENT OF THROMBOSIS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/965,099

; FILING DATE: 26-Sep-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/346,487  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: <Unknown>  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 73:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 1...321  
; OTHER INFORMATION: F9HZLC 2-0  
; SEQUENCE DESCRIPTION: SEQ ID NO: 73:  
US-09-965-099-73

Query Match 16.7%; Score 53; DB 10; Length 321;  
Best Local Similarity 100.0%; Pred. No. 3.8e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCCTGGTACCAGCAGAGCCAGG 119  
|||||  
Db 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCCTGGTACCAGCAGAGCCAGG 119

RESULT 14  
US-10-051-852-73  
; Sequence 73, Application US/10051852  
; Patent No. US20020146411A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Church, William  
; Gross, Mitchell  
; Feuerstein, Giora  
; Nichols, Andrew  
; Padlan, Eduardo  
; Patel, Arunbhai  
; Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; OF THROMBOSIS  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/051,852  
; FILING DATE: 17-Jan-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/344,050  
; FILING DATE: 25-JUN-1999  
; APPLICATION NUMBER: 08/783,853  
; FILING DATE: 16-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: <Unknown>  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 73:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 1...321  
; OTHER INFORMATION: F9HZLC 2-0  
; SEQUENCE DESCRIPTION: SEQ ID NO: 73:  
US-10-051-852-73

Query Match 16.7%; Score 53; DB 12; Length 321;  
Best Local Similarity 100.0%; Pred. No. 3.8e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCCTGGTACCAGCAGAGCCAGG 119  
|||||  
Db 67 TGCAGGCCAGCTCAAGTGTAAATTACATGCCTGGTACCAGCAGAGCCAGG 119

RESULT 15  
US-09-808-037-27  
; Sequence 27, Application US/09808037  
; Patent No. US20020052311A1  
; GENERAL INFORMATION:  
; APPLICANT: SOLOMON, Beka  
; APPLICANT: HANAN, Eliat  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF  
; FILE REFERENCE: NEUROLOGICAL DISEASES AND DISORDERS  
; FILE REFERENCE: SOLOMON-2D  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 09/629,971  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 09/473,653  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: US 60/152,417  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 27  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: synthetic construct  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(357)  
US-09-808-037-27

Query Match 16.7%; Score 53; DB 10; Length 357;  
Best Local Similarity 100.0%; Pred. No. 3.9e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 157 CTGGCTTCTCGAGTCCCTCGCTTCAGTGGCAGGAGGGTCTCGGACCTCTTA 209

Db 196 CTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGTCTCTGGGACCTTTA 248  
|||||

Search completed: November 27, 2002, 07:14:52  
Job time : 36.6419 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:19:05 ; Search time 37.5525 Seconds  
(without alignments)  
375.129 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 106

Sequence: 1 QIVLSQSPAILASPGKVT.....CQWSSNPPTFGGTMLEIR 106

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	44.3	106	AAW24532	Anti-Factor IX MAB
2	47	44.3	106	AAU81002	Mouse-human light
3	47	44.3	107	AAW24520	Mouse anti-human F
4	47	44.3	107	AAU80976	Murine BC2 light c
5	47	44.3	112	AAW24531	Anti-Factor IX MAB
6	47	44.3	112	AAU81001	Murine BC2 light c
7	46	43.4	105	AAW19914	Kappa light chain
8	43	40.6	633	AA1984965	Amino acid sequenc
9	42	39.6	106	AAU72834	Anti-NKG2D hybrido
10	42	39.6	108	AAU79883	Anti-EGFR antibody

11	42	39.6	412	22	AAW30694	A fusion of anti-C
12	42	39.6	423	22	AAW30695	A fusion of anti-C
13	41	38.7	128	8	AAW70628	Sequence encoded b
14	41	38.7	128	10	AAW94781	2 H7 VL gene. Syn
15	41	38.7	128	18	AAW10589	2H7 antibody light
16	41	38.7	128	18	AAW16344	2H7 light chain va
17	41	38.7	128	18	AAW10243	Light chain variab
18	41	38.7	128	19	AAW17521	Mouse 2H7 antibody
19	41	38.7	128	19	AAW41071	Mouse 2H7 antibody
20	41	38.7	128	19	AAW47514	Mouse 2H7 antibody
21	41	38.7	128	20	AAW89541	Mouse antibody 2H7
22	41	38.7	128	22	AAW89093	2H7 light chain va
23	40	37.7	129	13	AAW20789	Unprocessed variab
24	40	37.7	129	13	AAW30430	ASB7 antibody ligh
25	40	37.7	235	17	AAW06178	Murine ASB57 Light
26	40	37.7	235	20	AAW82740	Plasmid pNG3/ASB7
27	40	37.7	235	20	AAW82746	Plasmid pEE14/ASB7
28	35	33.0	145	12	AAW15323	IL-2 chimeric anti
29	35	33.0	145	14	AAW32127	Anti-IL2R-alpha an
30	32	30.2	32	20	AAW94743	Anti-Staph (HAV) 9
31	31	29.2	128	15	AAW55214	Murine variable re
32	31	29.2	235	21	AAW08025	A dimeric anti-CD2
33	27	25.5	105	17	AAW03182	Guy's 13 anti-Stre
34	27	25.5	106	13	AAW24105	Anti-Tac light cha
35	27	25.5	106	19	AAW58511	Protein SEQ ID NO:
36	27	25.5	106	22	AAE13144	Humanised antibody
37	27	25.5	106	22	AAE13144	Murine anti-Tac an
38	27	25.5	106	22	AAW69641	Murine anti-Tac an
39	27	25.5	106	23	AAW69651	Antigen-binding pr
40	27	25.5	107	14	AAU74418	Mab32 V-gene heavy
41	27	25.5	107	22	AAW83159	Mouse ganglioside
42	27	25.5	107	22	AAW83167	Ganglioside GM2 an
43	27	25.5	108	13	AAW21294	Murine VL kappa gr
44	27	25.5	108	16	AAW79884	Anti-EGFR antibody
45	27	25.5	108	18	AAW41388	Anti-CEA antibody

## ALIGNMENTS

## RESULT 1

AAW24532  
ID AAW24532 standard; Protein; 106 AA.

XX AAW24532;

XX 27-DEC-1997 (first entry)

XX Anti-Factor IX MAB chimeric light chain.

XX Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;  
chimeric antibody; antibody engineering; light chain.

OS Chimeric Mus musculus.

OS Chimeric Homo sapiens.

XX Chimeric synthetic.

PN WO9726010-A1.

PD 24-JUL-1997.

PF 17-JAN-1997; 97WO-US000759.

PR 24-OCT-1996; 96US-0029119.

PR 17-JAN-1996; 96US-0010108.

PA (SMIX ) SMITHKLINE BEECHAM CORP.

PA (UVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

XX Blackburn MN, Church WR, Feuerstein GZ, Gross MS;

PI Nichols AJ, Padlan EA, Patel AH, Sylvester DR;

XX WPI; 1997-385117/35.



CC This sequence comprises the light chain variable region of mouse  
 CC anti-human Factor IX monoclonal antibody BC2. Claimed humanised  
 CC antibodies (see AAW24510-18) contain CDRs (see AAW24504-09) of BC2  
 CC heavy and light chain variable regions inserted into framework  
 CC regions of selected human antibody sequences. They have self-  
 CC limiting neutralising activity, and are useful as anticoagulant  
 CC agents in treatment of thrombosis associated with myocardial  
 CC infarction, unstable angina, atrial fibrillation, stroke, renal  
 CC damage, pulmonary embolism, deep vein thrombosis, percutaneous  
 CC transluminal coronary angioplasty, disseminated intravascular  
 CC coagulation, sepsis, or artificial organs, shunts or prostheses  
 CC (claimed). Also claimed are chimeric antibodies (see AAW24532), and  
 CC Fab and Fab'/2 fragments. The claimed antibodies do not cause  
 CC uncontrolled bleeding (contrast heparin and warfarin) since they  
 CC provide only partial inhibition of coagulation.

XX Sequence 107 AA;

Query Match 44.3%; Score 47; DB 18; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-35;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSOSPAILSPGKVTMTCRASSVNYMHYQKPGSSPKPWI 47  
 |||||  
 Db 1 QIVLSOSPAILSPGKVTMTCRASSVNYMHYQKPGSSPKPWI 47

#### RESULT 4

AAU80976  
 ID AAU80976 standard; Protein; 107 AA.

XX AC AAU80976;

XX 09-APR-2002 (first entry)

DE Murine BC2 light chain variable region.

XX Human; mouse; BC2; animal post-thromboembolic induced ischaemia;  
 KW thrombolytic agent; anti-factor IX antibody; plasminogen activator;  
 KW thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;  
 KW vasotropic; cardiant; anti-respiratory syncytial virus;  
 KW heavy chain variable region; light chain variable region.

XX Mus sp.

XX WO200187339-A1.

XX 22-NOV-2001.

XX 05-OCT-2000; 2000WO-US27438.

XX 15-MAY-2000; 2000US-0571434.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;

XX WPI; 2002-082944/11.

XX N-PSDB; ABK23937.

XX Treating post-thromboembolic induced ischaemia in an animal by  
 PT administering anti-factor IX antibody in combination with a plasminogen  
 PT activator

XX Example 5; Page 97; 163pp; English.

XX The invention relates to a method for treating an animal  
 CC post-thromboembolic induced ischaemia or reducing a required dose of a  
 CC thrombolytic agent in treatment of an animal post-thromboembolic induced  
 CC ischaemia, comprising administering an anti-factor IX antibody or its  
 CC fragment, optionally in combination with a plasminogen activator or  
 CC thrombolytic agent. The method is useful for treating  
 CC post-thromboembolic-induced ischaemia, for preventing thromboembolic

CC stroke in an animal, and for reducing a required dose of a thrombolytic  
 CC agent. Sequences AAU80972-AAU81004 represent antibodies and vector  
 CC polypeptides used in the method of the invention.

XX Sequence 107 AA;

Query Match 44.3%; Score 47; DB 23; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-35;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSOSPAILSPGKVTMTCRASSVNYMHYQKPGSSPKPWI 47  
 |||||  
 Db 1 QIVLSOSPAILSPGKVTMTCRASSVNYMHYQKPGSSPKPWI 47

#### RESULT 5

AAW24531  
 ID AAW24531 standard; Protein; 112 AA.

XX AC AAW24531;

XX 27-DEC-1997 (first entry)

XX Anti-Factor IX MAB BC2 modified light chain.

XX Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;  
 KW chimeric antibody; antibody engineering; light chain.

XX Chimeric Mus musculus.

XX Chimeric synthetic.

XX WO9726010-A1.

XX 24-JUL-1997.

XX 17-JAN-1997; 97WO-US00759.

XX 24-OCT-1996; 96US-0029119.

XX 17-JAN-1996; 96US-0010108.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

XX Blackburn MN, Church WR, Feuerstein GZ, Gross MS;

XX Nichols AJ, Padlan EA, Patel AH, Sylvester DR;

XX WPI; 1997-385117/35.

XX N-PSDB; AAT79899.

XX Inhibiting thrombosis with self-limiting antibody to coagulation  
 PT factor - avoids uncontrolled bleeding by providing only partial  
 PT inhibition

XX Example 7; Page 126; 150pp; English.

XX This polypeptide sequence comprises a modified light chain variable  
 CC region (see also AAW24520) of mouse anti-human factor IX monoclonal  
 CC antibody BC2. It is encoded by a DNA fragment (see AAT79899)  
 CC produced by PCR amplification of BC2 cDNA. The modified BC2 VL  
 CC region is incorporated into novel mouse-human chimeric light chain  
 CC F9CHLC (see AAW24532). Claimed anti-Factor IX chimeric antibodies  
 CC are useful in the treatment of thrombosis.

XX Sequence 112 AA;

Query Match 44.3%; Score 47; DB 18; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-35;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSOSPAILSPGKVTMTCRASSVNYMHYQKPGSSPKPWI 47  
 |||||  
 Db 1 QIVLSOSPAILSPGKVTMTCRASSVNYMHYQKPGSSPKPWI 47

RESULT 6  
AAU81001  
ID AAU81001 standard; Protein: 112 AA.  
XX  
AC AAU81001;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Murine BC2 light chain modified variable region.  
XX  
KW Human; mouse; BC2; animal post-thromboembolic induced ischaemia;  
KW thrombolytic agent; anti-factor IX antibody; plasminogen activator;  
KW thromboembolic stroke; cerebroprotective; anticoagulant; Thrombolytic;  
KW vasotropic; cardiant; anti-respiratory syncytial virus;  
KW heavy chain variable region; light chain variable region.  
XX  
OS Mus sp.  
OS Synthetic.  
XX  
XX WO200187339-A1.  
XX  
XX 22-NOV-2001.  
XX  
XX 05-OCT-2000; 2000WO-US27438.  
XX  
XX 15-MAY-2000; 2000US-0571434.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX  
XX Barone FC, Blackburn MN, Feuerstein GZ, Toomey JR;  
XX  
XX WPI; 2002-082944/11.  
XX  
XX N-PSDB; ABK24004.  
XX  
XX Treating post-thromboembolic induced ischaemia in an animal by  
XX administering anti-factor IX antibody in combination with a plasminogen  
XX activator -  
XX  
XX Example 7; Page 153-154; 163pp; English.  
XX  
XX The invention relates to a method for treating an animal  
XX post-thromboembolic induced ischaemia or reducing a required dose of a  
XX thrombolytic agent in treatment of an animal post-thromboembolic induced  
XX ischaemia, comprising administering an anti-factor IX antibody or its  
XX fragment, optionally in combination with a plasminogen activator or  
XX thrombolytic agent. The method is useful for treating  
XX post-thromboembolic-induced ischaemia, for preventing thromboembolic  
XX stroke in an animal, and for reducing a required dose of a thrombolytic  
XX agent. Sequences AAU80972-AAU81004 represent antibodies and vector  
XX polypeptides used in the method of the invention.  
XX  
SQ Sequence 112 AA;  
  
Query Match 44.3%; Score 47; DB 23; Length 112;  
Best Local Similarity 100.0%; Pred. No. 7.7e-35;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QIVLSQSPAILASPGKVTMTCRASSSVNMYHMYQKPGSSPKPWI 47  
DB 1 QIVLSQSPAILASPGKVTMTCRASSSVNMYHMYQKPGSSPKPWI 47  
  
RESULT 7  
AAU19914  
ID AAU19914 standard; Protein: 105 AA.  
XX  
AC AAU19914;  
XX  
DT 24-FEB-1998 (first entry)  
XX  
XX Kappa light chain of MAb 5C3.

TrkA receptor; monoclonal antibody; docking site; human p140 TrkA; NGF;  
hormone-receptor interactive site mapping; artificial receptor agonist;  
receptor-domain functional correlation; NGF docking site; therapy;  
nerve growth factor; tumour growth inhibitor; neurological disease;  
neuroma; neoplastic tumour.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
XX 24..33  
FT Region /note= "complementarity determining region 1"  
FT Region 49..55  
FT Region /note= "complementarity determining region 2"  
FT Region 88..96  
FT Region /note= "complementarity determining region 3"  
XX  
XX WO9721732-A1.  
XX  
XX 19-JUN-1997.  
XX  
XX 06-DEC-1996; 96WO-CA00815.  
XX  
XX 08-DEC-1995; 95GB-0025180.  
XX  
XX (OYMC-) UNIV MCGILL.  
XX  
XX Lesauter L, Saragovi HU;  
XX  
XX WPI; 1997-332727/30.  
XX  
XX Anti-human TrkA antibody recognising nerve growth factor docking  
XX site - useful e.g. for treatment of neurological diseases,  
XX neuroma(s) or neoplastic tumours  
XX  
XX Example 3; Page 31; 61pp; English.  
XX  
XX AAU19914 and AAU19915 represent the kappa light chain and the heavy chain  
XX of the monoclonal antibody (MAb) 5C3. MAb 5C3 is an example of an MAb of  
XX the invention. The MAb of the invention, optionally having  
XX complementarity determining region (CDR)-like domains of hormones, binds  
XX the TrkA receptor under physiological conditions. MAb 5C3, directed  
XX against human p140 TrkA, is a structural and functional mimic of NGF and  
XX an artificial receptor agonist. It binds in the NGF docking site and  
XX promotes TrkA internalisation, TrkA and phosphotyrosine kinase  
XX tyrosine phosphorylation and increases transformation of TrkA-expressing  
XX fibroblasts. The Ab or a TrkA docking site, are used to screen for  
XX agents that inhibit or mimic nerve growth factor (NGF) biological  
XX activity. The agents and Ab inhibit tumour growth in situ for treatment  
XX or prevention of neurological diseases, neuromas and neoplastic tumours  
XX expressing TrkA receptor by preventing binding of NGF to TrkA receptor or  
XX down-regulating the receptor. The Ab is also used for mapping  
XX hormone-receptor interactive sites and receptor-domain functional  
XX correlation. The Ab can also induce NGF-like responses, especially  
XX differentiation/neurogenesis or trophic functions, by cells expressing  
XX TrkA. The prognosis and diagnosis of human tumours can be performed using  
XX the Ab in biopsies and immunocytochemistry or by labelling the Ab and  
XX nuclear imaging. Treatment of these tumours involves coupling the Ab to a  
XX cytotoxic agent and administering the complex to the patient.  
XX  
SQ Sequence 105 AA;  
  
Query Match 43.4%; Score 46; DB 18; Length 105;  
Best Local Similarity 100.0%; Pred. No. 5.8e-34;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 49 ATSNLASGVPARFSGSGSGTSLTISRVEAEDAATYYCQWSSNP 94  
DB 49 ATSNLASGVPARFSGSGSGTSLTISRVEAEDAATYYCQWSSNP 94  
  
RESULT 8  
AAU84965  
ID AAU84965 standard; Protein: 633 AA.



```

XX AAY84965;
XX
XX 21-AUG-2000 (first entry)
XX
XX Amino acid sequence of a CD-20 specific chimeric receptor.
XX
XX CD20-specific receptor; CD-20 specific redirected T cell; leukemia;
XX CD20+ malignancy; non-Hodgkin's lymphoma; myeloablative chemotherapy;
XX stem cell rescue; autoimmune disease; lupus; rheumatoid arthritis.
XX
XX Synthetic.
OS Mus sp.
XX
XX Key Location/Qualifiers
XX Peptide 27..86
XX /note= "signal peptide from murine T86.66 antibody
XX kappa light chain"
XX
XX Region 21..126
XX /note= "anti-CD20 variable regions"
XX Peptide 145..266
XX /note= "GS18 linker"
XX Region 283..392
XX /note= "hinge region"
XX Region 393..499
XX /note= "CH3 region"
XX Region 500..521
XX /note= "CD4 transmembrane region"
XX Region 522..633
XX /note= "zeta chain"
XX
XX WO200023573-A2.
XX
XX 27-APR-2000.
XX
XX 20-OCT-1999; 99WO-US24484.
XX
XX 20-OCT-1998; 98US-0105014.
XX
XX (CITY ) CITY OF HOPE.
XX
XX Raubitschek A, Jensen MC, Wu AM;
XX
XX WPI; 2000-339676/29.
XX
XX N-PSDB; AAA15019.
XX
XX Genetically engineered CD20-specific redirected T cells useful for
XX treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+
XX acute or chronic leukemia, and autoimmune disease -
XX
XX Claim 10; Page 53-55; 58pp; English.
XX
XX The present sequence represents a synthetic CD20-specific chimeric
XX receptor. The specification describes CD-20 specific redirected T cells
XX which express and bear on the cell surface membrane a CD20-chimeric
XX receptor comprising an intracellular signalling domain, a transmembrane
XX domain and an extracellular domain, the extracellular domain comprising
XX a CD20-specific receptor. The genetically engineered CD20-specific
XX redirected T cells are useful for treating a CD20+ malignancy, such
XX as non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a
XX human patient having previously undergone myeloablative chemotherapy and
XX stem cell rescue. The genetically engineered CD20-specific redirected
XX T cells are also useful for abrogating an untoward B cell function, such
XX as autoimmune disease (lupus or rheumatoid arthritis) in a patient.
XX
XX Sequence 633 AA:
XX
XX Query Match 40.6%; Score 43; DB 21; Length 633;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-30;
XX Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 49 ATSNLASGVPARFSGSGTSLTISRVEAEDAATYYCQOWS 91
XX |||||||||||||||||||||||||||||||||||||||||||

```

```

Db 69 ATSNLASGVPARFSGSGTSLTISRVEAEDAATYYCQOWS 111
XX
XX RESULT 9
XX AAU72834
XX
XX AAU72834 standard; Peptide; 106 AA.
XX
XX AAU72834;
XX
XX 26-FEB-2002 (first entry)
XX
XX Anti-NKG2D hybridoma 6H7E7 variable light chain CDRL1.
XX
XX Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;
XX autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;
XX intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;
XX prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;
XX sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAPI10;
XX helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;
XX 8G7C10; 6F5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;
XX P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;
XX p53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.
XX
XX Homo sapiens.
XX
XX WO200171005-A2.
XX
XX 27-SEP-2001.
XX
XX 26-MAR-2001; 2001WO-EP03414.
XX
XX 24-MAR-2000; 2000EP-0106467.
XX
XX (KUFE/) KUFER P.
XX
XX Kufer P, Riethmueller G, Lutterbuese R, Borschert K, Kischel R;
XX Mayer M, Hofmeister R;
XX
XX WPI; 2002-055119/07.
XX
XX Multifunctional polypeptides comprising binding sites that specifically
XX recognise extracellular groups of the NKG2D receptor complex and
XX domains which function as receptors or ligands, useful for treating
XX cancers and infectious diseases -
XX
XX Example 3; Fig 16; 114pp; English.
XX
XX The invention relates to a multifunctional polypeptide comprising a
XX domain with a binding site that specifically recognises an extracellular
XX group of the NKG2D receptor complex and a second domain which functions
XX as a receptor or ligand. The polypeptide and its associated
XX polynucleotide are used for the preparation of a pharmaceutical
XX composition for the treatment of cancer, infections and/or autoimmune
XX conditions. The cancer may be a tumour of the head and neck, stomach,
XX oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,
XX larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,
XX bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.
XX The infectious diseases can be caused by viruses, bacteria, fungi,
XX protozoa or helminths. The autoimmune diseases include multiple
XX sclerosis, Grave's disease, ankylosing spondylitis, acute anterior
XX uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent
XX diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and
XX autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKG2D
XX receptor and the polypeptides of the invention.
XX
XX Sequence 106 AA:
XX
XX Query Match 39.6%; Score 42; DB 23; Length 106;
XX Best Local Similarity 100.0%; Pred. No. 2.4e-30;
XX Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 49 ATSNLASGVPARFSGSGTSLTISRVEAEDAATYYCQOW 90
XX |||||||||||||||||||||||||||||||||||||||||||

```

```
Db 49 ATSNLASGVPAFSGSGTSTSLTISRVEADAATYYCQW 90
RESULT 10
AAR79883
ID AAR79883 standard; Protein; 108 AA.
XX
AC AAR79883;
XX
DT 02-JUL-1996 (first entry)
XX
DE Anti-EGFR antibody light chain variable region (Clone L2 1C).
XX
KW Single chain antibody; antibody; epidermal growth factor receptor;
KW EGFR; tumour; cancer; glioma; melanoma; carcinoma; diagnosis;
KW assessment; phage antibody library.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 1..23
FT /label= Framework region 1.
FT Region 24..33
FT /label= CDR1.
FT Region 34..48
FT /label= Framework region 2.
FT Region 49..55
FT /label= CDR2.
FT Region 56..87
FT /label= Framework region 3.
FT Region 88..96
FT /label= CDR3.
FT Region 97..108
FT /label= Framework region 4.
XX
WO9525167-A1.
XX
PN 21-SEP-1995.
XX
PD 16-MAR-1995; 95WO-EP00978.
XX
PE 02-DEC-1994; 94EP-0118970.
XX
PR 17-MAR-1994; 94EP-0104160.
XX
PA (MERE ) MERCK PATENT GMBH.
XX
PI Adan J, Ansell KH, Bendig MM, Blasco F, Guessow D;
PI Kettleborough AC, Mitjans F, Piulats J, Rosell E;
XX
WPI; 1995-336972/43.
XX
PT Anti-EGFR antibodies and single chain Fv antibody fragments -
PT obtained from phage-antibody libraries, useful for diagnosis and
PT therapy of tumours
XX
PS Disclosure; Figure 1A; 93pp; English.
XX
CC Anti-epidermal growth factor receptor (EGFR) single chain antibodies
CC and antibodies constructed from anti-EGFR antibody fragments can be
CC used for diagnosis of tumours and assessment of tumour growth in
CC vitro and in vivo. They may also be used in a pharmaceutical
CC composition for the therapy of e.g. melanomas, gliomas and carcinomas.
CC The antibodies and fragments are derived from mice but are humanised
CC so as to cause minimum reaction against them. They are produced
CC using the phage antibody library.
XX
SQ Sequence 108 AA;
Query Match 39.6%; Score 42; DB 16; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 49 ATSNLASGVPAFSGSGTSTSLTISRVEADAATYYCQW 90
Db 49 ATSNLASGVPAFSGSGTSTSLTISRVEADAATYYCQW 90
RESULT 11
AAB30694
ID AAB30694 standard; Protein; 412 AA.
XX
AC AAB30694;
XX
DT 02-APR-2001 (first entry)
XX
DE A fusion of anti-CD20 single chain antibody/streptavidin.
XX
KW Streptavidin; tumour cell; cancer; adenocarcinoma;
KW hematological malignancy; B9E9.
XX
OS Synthetic.
OS Streptomyces avidinii.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..108
FT /note= "VL chain"
FT Peptide 109..126
FT /note= "linker"
FT Protein 127..248
FT /note= "VH chain"
FT Peptide 249..253
FT /note= "linker"
FT Protein 254..412
FT /note= "streptavidin"
XX
WO200075333-A1.
XX
PN 14-DEC-2000.
XX
PD 05-JUN-2000; 2000WO-US15595.
XX
PE 07-JUN-1999; 99US-0137900.
XX
PR 03-DEC-1999; 99US-0168976.
XX
PA (NEOR-) NEORX CORP.
XX
PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;
XX
WPI; 2001-091213/10.
XX
N-PSDB; AAC86563.
XX
PT New vector constructs for expressing genomic streptavidin fusion
PT proteins which are useful for targeting tumour cells associated with
PT cancer, e.g. adenocarcinomas -
XX
PS Example 2; Fig 11B; 100pp; English.
XX
CC The present sequence represents a fusion of an anti-CD20 single chain
CC antibody (B9E9) streptavidin. The fusion protein is expressed using
CC vectors of the invention. The specification describes vector constructs
CC for expressing streptavidin fusion proteins. The vector comprises a
CC nucleic acid encoding streptavidin or its functional variant operatively
CC linked to a promoter, and a cloning site for insertion of a second
CC nucleic acid sequence encoding a polypeptide to be fused with
CC streptavidin, interposed between the promoter and the first nucleic
CC acid sequence. Alternatively, the vector construct comprises a nucleic
CC acid, operatively linked to a promoter, encoding a polypeptide to be
CC fused with streptavidin, and a cloning site for insertion of a second
CC nucleic acid encoding at least 129 amino acids of streptavidin or its
CC functional variant. The fusion proteins are useful for targeting tumour
CC cells, particularly tumour cells associated with cancer,
CC e.g. adenocarcinomas or hematological malignancies. The vector construct
CC is useful for expressing of streptavidin fusion proteins. In particular,
CC these are useful as tools for medical diagnostics and therapeutic
CC purposes, e.g. for detecting the presence or absence of, or treating, a
```

CC target site within a mammalian host.

XX Sequence 412 AA;

Query Match 39.6%; Score 42; DB 22; Length 412;

Best Local Similarity 100.0%; Pred. No. 7.8e-30;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQOW 90

Db 49 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQOW 90

RESULT 12

AAB30695

ID AAB30695 standard; Protein: 423 AA;

XX AC AAB30695;

XX DT 02-APR-2001 (first entry)

XX DE A fusion of anti-CD20 single chain antibody/streptavidin.

XX KW Streptavidin; tumour cell; cancer; adenocarcinoma;

XX OS hematological malignancy; B9E9.

XX OS Synthetic.

XX OS Streptomyces avidinii.

XX OS Homo sapiens.

FH Key Location/Qualifiers

FT Protein 1..126

FT Peptide /note= "VH chain"

FT Peptide 127..150

FT Peptide /note= "linker"

FT Protein 151..259

FT Peptide /note= "VL chain"

FT Peptide 260..264

FT Protein /note= "linker"

FT Protein 265..423

FT Protein /note= "streptavidin"

XX WO200075333-A1.

XX 14-DEC-2000.

XX 05-JUN-2000; 2000WO-US15595.

XX 07-JUN-1999; 99US-0137900.

XX 03-DEC-1999; 99US-0168976.

XX (NEOR-) NEORX CORP.

XX PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;

XX WPI; 2001-091213/10.

XX N-PSDB; AAC86564.

XX New vector constructs for expressing genomic streptavidin fusion

XX proteins which are useful for targeting tumour cells associated with

XX cancer, e.g. adenocarcinomas -

XX Example 2; Fig 11C; 100pp; English.

XX The present sequence represents a fusion of an anti-CD20 single chain  
CC antibody (B9E9) streptavidin. The fusion protein is expressed using  
CC vectors of the invention. The specification describes vector constructs  
CC for expressing streptavidin fusion proteins. The vector comprises a  
CC nucleic acid encoding streptavidin or its functional variant operatively  
CC linked to a promoter, and a cloning site for insertion of a second  
CC nucleic acid sequence encoding a polypeptide to be fused with  
CC streptavidin, interposed between the promoter and the first nucleic  
CC acid sequence. Alternatively, the vector construct comprises a nucleic

CC acid, operatively linked to a promoter, encoding a polypeptide to be  
CC fused with streptavidin, and a cloning site for insertion of a second  
CC nucleic acid encoding at least 129 amino acids of streptavidin or its  
CC functional variant. The fusion proteins are useful for targeting tumour  
CC cells, particularly tumour cells associated with cancer,  
CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
CC is useful for expressing of streptavidin fusion proteins. In particular,  
CC these are useful as tools for medical diagnostics and therapeutic  
CC purposes, e.g. for detecting the presence or absence of, or treating, a  
CC target site within a mammalian host.

XX Sequence 423 AA;

Query Match

Best Local Similarity 39.6%; Score 42; DB 22; Length 423;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQOW 90

Db 200 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQOW 241

RESULT 13

AAP70628

ID AAP70628 standard; protein: 128 AA.

XX AC AAP70628;

XX DT 03-OCT-2002 (updated)

XX DT 09-APR-1991 (first entry)

XX DE Sequence encoded by the 2H7 VL sequence in which the Vk gene

XX DE contains JK5 sequences.

XX KW Chimeric antibody; Anti-cancer antibody.

XX OS Chimeric - Homo sapiens.

XX OS Chimeric - Mus sp.

FH Key Location/Qualifiers

FT Region 1..23

FT /tag= a

FT /label=leader peptide

FT Region 24..47

FT /tag= b

FT /label=FR1

FT Region 48..56

FT /tag= c

FT /label=CDR1

FT Region 57..71

FT /tag= d

FT /label=FR2

FT Region 72..78

FT /tag= e

FT /label=CDR2

FT Region 79..110

FT /tag= f

FT /label=FR3

FT Region 111..119

FT /tag= g

FT /label=CDR3

FT Region 118..128

FT /tag= h

FT /label=JK5

FT Region 120..128

FT /tag= i

FT /label=FR4

XX WO8702671-A.

XX 07-MAY-1987.

XX 27-OCT-1986; 86WO-US02269.

```
XX 01-NOV-1985; 85US-0793980.
XX (ITGE-) INT GENETIC ENG INC.
PA (ROBI/) ROBINSON R R.
XX Robinson RR, Liu AY, Horwitz AH, Wall R;
XX WPI; 1987-136004/19.
DR N-PSDB; AAN70972.
XX
PT Prodn. of immunoglobulin chains and molecules - is by recombinant
PT DNA procedures, with chimeric antibodies etc. related to cancer
PT specific antigens.
XX
PS Example; Fig 22; 126pp; English.
XX
CC The patentors claim a chimeric antibody molecule comprising 2 light
CC chains and 2 heavy chains, each comprising a constant human region
CC and a variable non-human region. Coding sequences for the Ig chains
CC are also claimed. The invention provides consensus sequences of
CC light and heavy chain J regions useful in the design of
CC oligonucleotides (UIGs) for use as primers or probes for cloning
CC immunoglobulinlight or heavy chain mRNAs or genes. Depending on the
CC nature of design of a particular UIG, it may be capable of
CC hybridizing to all Ig mRNAs or genes containing a single specific J
CC sequence. UIG denotes universal immunoglobulin gene.
CC (Updated on 03-OCT-2002 to add missing OS field.)
XX
SQ Sequence 128 AA;
Query Match 38.7%; Score 41; DB 8; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.3e-29;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 SNLASGVPARFSGSGTSLTISRVEADAATYYCQOWS 91
Db 73 SNLASGVPARFSGSGTSLTISRVEADAATYYCQOWS 113
|||||
RESULT 14
AAP94781
ID AAP94781 standard; protein; 128 AA.
XX
AC AAP94781;
XX
DT 06-JUL-1990 (first entry)
XX
DE 2 H7 VL gene.
XX
KW Antibodies; passive immunisation; pH3-12a; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label=leader peptide.
FT Domain 23..46
FT /label=FR1.
FT Domain 50..55
FT /label=CDR1
FT Domain 56..70
FT /label=FR2
FT Domain 71..77
FT /label=CDR2.
FT Domain 78..109
FT /label=FR3.
FT Domain 110..118
FT /label=CDR3.
FT Domain 119..128
FT /label=FR4.
XX
PN W08900999-A.
```

```
XX 09-FEB-1989.
XX
XX 25-JUL-1988; 88WO-US02514.
XX
XX 24-JUL-1987; 87US-0077528.
XX
XX (ITGE-) INT GENETIC ENG INC.
XX
XX Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;
XX
XX WPI; 1989-061144/08.
DR N-PSDB; AAN91147.
XX
XX Polynucleotide(s) encoding immunoglobulin molecules -
XX used for efficient prodn. of chimeric human or non-human or
XX class switched antibodies.
XX
XX Disclosure; 7pp; English.
XX
XX Sequence carries the variable region of the chimeric immunoglobulin
XX sequence. The antibodies are useful in passive immunisation avoiding
XX negative immune reactions. They are also useful in assaying and in vitro
XX imaging.
XX
XX Sequence 128 AA;
Query Match 38.7%; Score 41; DB 10; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.3e-29;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 SNLASGVPARFSGSGTSLTISRVEADAATYYCQOWS 91
Db 73 SNLASGVPARFSGSGTSLTISRVEADAATYYCQOWS 113
|||||
RESULT 15
AAW10589
ID AAW10589 standard; Protein; 128 AA.
XX
AC AAW10589;
XX
DT 21-OCT-1997 (first entry)
XX
DE 2H7 antibody light chain variable region.
XX
KW Immunoglobulin G; IgG; light chain; recombinant production;
KW antibody; passive immunisation; serum sickness; anaphylactic shock;
KW immunoassay; imaging; reagent; complement mediated lysis;
KW therapy; variable; region; complementarity determining; CDR;
KW framework.
XX
OS Mus spp.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= sig_peptide
FT Peptide 23..128
FT /label= mat_peptide
FT Region 23..46
FT /label= framework_1
FT Region 47..55
FT /label= CDR_1
FT Region 56..70
FT /label= framework_2
FT Region 71..77
FT /label= CDR_2
FT Region 78..109
FT /label= framework_3
FT Region 110..118
FT /label= CDR_3
FT Region 119..128
FT /label= framework_4
```

```
XX US5595898-A.
PN
XX
XX PD 21-JAN-1997.
XX
XX PF 01-NOV-1985; 85US-0793980.
XX
XX PR 29-MAR-1990; 90US-0501092.
XX PR 01-NOV-1985; 85US-0793980.
XX PR 27-OCT-1986; 86WO-US02289.
XX PR 24-JUL-1987; 87US-0077528.
XX PR 11-JAN-1988; 88US-0142039.
XX PR 08-DEC-1992; 92US-0987555.
XX PR 18-AUG-1994; 94US-0299085.
XX
XX PA (XOMA ) XOMA CORP.
XX
XX PI Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;
XX PI Wall R, Wilcox GL;
XX
XX DR WPI; 1997-107579/10.
XX DR N-PSDB; AAT36317.
XX
XX PT Nucleic acid encoding immunoglobulin fragment - comprising
XX PT di:cistronic transcription unit with pectate lyase signal sequences
XX
XX PS Example; Fig 22; 95pp; English.
XX
XX CC The present sequence is the light chain variable region of
XX CC the 2H7 antibody. The 2H7 cDNA was used in the preparation of
XX CC a novel polynucleotide molecule encoding an Ig fragment. The DNA
XX CC molecule comprises 2 DNA sequences encoding 2 pectate lyase
XX CC secretion signal sequences respectively linked to a DNA sequence
XX CC encoding an Ig Fd molecule or Ig light chain, operably linked to a
XX CC single prokaryotic promoter so as to form a dicistronic
XX CC transcription unit, provided that the Ig fragment can bind an
XX CC antigen and is produced and secreted by an E. coli host cell when
XX CC the nucleic acid molecule is expressed in the host cell.
XX CC The polynucleotide molecule is used for the production of
XX CC recombinant antibodies, which can be used for passive immunisation
XX CC without negative immune reactions (e.g. serum sickness and
XX CC anaphylactic shock), in labelled forms as immunoassay or imaging
XX CC reagents, in complement mediated lysis and for therapeutic
XX CC purposes when coupled to a toxin or other therapeutic agent.
XX
XX SQ Sequence 128 AA;

Query Match 38.7%; Score 41; DB 18; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.3e-29;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 SNLASGVPARFSGSGTYSLTISRVEADAATYYCOOWS 91
DB 73 SNLASGVPARFSGSGTYSLTISRVEADAATYYCOOWS 113

Search completed: November 27, 2002, 07:28:00
Job time : 38.6525 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:25:20 ; Search time 15.9614 Seconds  
(without alignments)  
638.431 Million cell updates/sec

Title: US-09-893-615-89  
Perfect score: 106  
Sequence: 1 QIVLSQSPAILSPGKVT.....CQWSSNPPTFGGTMLEIR 106

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	37.7	106	2	Ig kappa chain V r
2	38	35.8	107	2	Ig kappa chain V r
3	29	27.4	107	2	Ig kappa chain V r
4	27	25.5	91	2	Ig kappa chain V r
5	27	25.5	93	2	Ig kappa chain V r
6	27	25.5	99	2	Ig kappa chain V r
7	26	24.5	86	2	Ig heavy chain V r
8	26	24.5	86	2	Ig heavy chain V r
9	26	24.5	90	2	Ig kappa chain V r
10	26	24.5	93	2	Ig kappa chain V r
11	26	24.5	93	2	Ig kappa chain V r
12	26	24.5	93	2	Ig kappa chain V r
13	26	24.5	94	2	Ig light chain V r
14	26	24.5	96	2	Ig kappa chain V r
15	26	24.5	97	2	Ig light chain V r
16	26	24.5	98	2	Ig light chain V r
17	26	24.5	99	2	Ig light chain V r
18	26	24.5	99	2	Ig light chain V r
19	26	24.5	105	2	Ig kappa chain V r
20	26	24.5	108	2	Ig kappa chain V r
21	26	24.5	108	2	Ig kappa chain V r
22	26	24.5	108	2	Ig kappa chain V r
23	26	24.5	129	1	Ig kappa chain pre
24	26	24.5	130	2	Ig kappa chain pre
25	26	24.5	130	2	Ig kappa chain pre
26	25	23.6	93	2	Ig kappa chain V r
27	25	23.6	93	2	Ig kappa chain V r
28	25	23.6	93	2	Ig kappa chain V r
29	25	23.6	93	2	Ig kappa chain V r

30	25	23.6	93	2	SI7634	Ig kappa chain V r
31	25	23.6	93	2	SI7635	Ig kappa chain V r
32	25	23.6	93	2	SI7636	Ig kappa chain V r
33	25	23.6	101	2	SI3696	Ig heavy chain V r
34	25	23.6	102	2	SI3697	Ig heavy chain V r
35	25	23.6	102	2	S29582	Ig kappa chain V r
36	25	23.6	102	2	S29588	Ig kappa chain V r
37	25	23.6	103	2	SI3695	Ig kappa chain - m
38	25	23.6	103	2	SI3698	Ig heavy chain V r
39	25	23.6	106	2	S29583	Ig kappa chain V r
40	25	23.6	108	2	PS0069	Ig kappa chain V r
41	25	23.6	109	2	SI3699	Ig kappa chain V r
42	25	23.6	109	2	PT0405	Ig light chain V r
43	25	23.6	120	2	S66536	Ig light chain V r
44	25	23.6	123	2	S05269	Ig kappa chain pre
45	25	23.6	124	2	S05267	Ig kappa chain pre

ALIGNMENTS

RESULT 1

PL0082

Ig kappa chain V region (2D3) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 21-Jan-2000  
C:Accession: PL0082  
R:Meek, K.; Hasemann, C.; Pollok, B.; Alkan, S.S.; Brait, M.; Slouli, M.; Urbain, J.; J. Exp. Med. 169, 519-533, 1989  
A:Title: Structural characterization of antidiotypic antibodies: evidence that Ab2s  
A:Reference number: PL0080; MUID:89094248; PMID:2492056  
A:Accession: PL0082  
A:Molecule type: mRNA  
A:Residues: 1-106 <MEM>  
A:Experimental source: strain BALB/c

A:Note: the sequence shown here is from the V kappa region of an antidiotypic monoclonal antibody  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:16-89/Domain: immunoglobulin homology <IMM>

Query Match 37.7%; Score 40; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 7.3e-33;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 SATSNLASGVPARFSGSGTSTSLTISRVEADAATTC 87  
Db 48 SATSNLASGVPARFSGSGTSTSLTISRVEADAATTC 87

RESULT 2

PC4405

Ig kappa chain V region (F3, anti-AFP) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Nov-1997 #sequence\_revision 23-Jan-1998 #text\_change 21-Jan-2000  
C:Accession: PC4405  
R:Deng, J.B.; Han, H.; Su, C.-Z.; Chen, C.O.  
Chinese Biochem. J. 12, 648-653, 1996

A:Title: Generation of a phage display library of the immunoglobulin repertoire from  
A:Reference number: PC4405  
A:Accession: PC4405  
A:Molecule type: mRNA  
A:Residues: 1-107 <DEN>  
A:Experimental source: spleen cell  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:16-89/Domain: immunoglobulin homology <IMM>

Query Match 35.8%; Score 38; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 7.6e-31;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 LASGVPARFSGSGTSTSLTISRVEADAATTCQW 90  
Db 53 LASGVPARFSGSGTSTSLTISRVEADAATTCQW 90

## RESULT 3

PD0011  
Ig kappa chain V region (VLB10, anti-AFP) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 21-Jan-2000  
C:Accession: PD0011  
R:Deng, J.B.; Han, H.; Su, C.-Z.; Chen, C.Q.  
Chinese Biochem J. 12, 648-653, 1996  
A:Title: Generation of a phage display library of the immunoglobulin repertoire from human  
A:Reference number: PC4405  
A:Contents: Spleen  
A:Accession: PD0011  
A:Molecule type: mRNA  
A:Residues: 1-107 <EN>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:8-89/Domain: immunoglobulin homology <IMM>

Query Match 27.4%; Score 29; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 8.8e-22;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 SSGSGTSLTISRVEADAATYYCQQW 90  
|||||  
Db 62 SSGSGTSLTISRVEADAATYYCQQW 90

## RESULT 4

SI7630  
Ig kappa chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C:Accession: SI7630  
R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.  
Nature 352, 624-628, 1991  
A:Title: Making antibody fragments using phage display libraries.  
A:Reference number: SI7230; MUID:91326098; PMID:1907718  
A:Accession: SI7630  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-91 <CLA>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:8-81/Domain: immunoglobulin homology <IMM>

Query Match 25.5%; Score 27; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 7.9e-20;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASGVPARFSGSGTSLTISR 76  
|||||  
Db 42 TSNLASGVPARFSGSGTSLTISR 68

## RESULT 5

SI7641  
Ig kappa chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C:Accession: SI7641  
R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.  
Nature 352, 624-628, 1991  
A:Title: Making antibody fragments using phage display libraries.  
A:Reference number: SI7230; MUID:91326098; PMID:1907718  
A:Accession: SI7641  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-93 <CLA>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:8-83/Domain: immunoglobulin homology <IMM>

Query Match 25.5%; Score 27; DB 2; Length 93;  
Best Local Similarity 100.0%; Pred. No. 8.1e-20;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASGVPARFSGSGTSLTISR 76  
|||||  
Db 44 TSNLASGVPARFSGSGTSLTISR 70

## RESULT 6

D38601  
Ig kappa chain V region (6A7) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Aug-1991 #sequence\_revision 30-Aug-1991 #text\_change 21-Jan-2000  
C:Accession: D38601  
R:Coshorn, S.C.; Retzel, E.; Jemerson, R.  
J. Biol. Chem. 266, 2134-2142, 1991  
A:Title: Common structural features among monoclonal antibodies binding the same anti  
A:Reference number: A38601; MUID:91115823; PMID:1703527  
A:Accession: D38601  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-99 <GOS>  
C:Cross-references: GB:M57981; NID:gl96408; PIDN:AAA63362.1; PID:gl96409  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:8-81/Domain: immunoglobulin homology <IMM>

Query Match 25.5%; Score 27; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 8.5e-20;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASGVPARFSGSGTSLTISR 76  
|||||  
Db 42 TSNLASGVPARFSGSGTSLTISR 68

## RESULT 7

S20649  
Ig heavy chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C:Accession: S20649  
R:Losman, M.; Fasy, T.M.; Novick, K.E.; Monestier, M.  
Submitted to the EMBL Data Library, February 1992  
A:Description: Relationships among antinuclear antibodies from autoimmune MRL mice re  
A:Reference number: S20639  
A:Accession: S20649  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-86 <LOS>  
C:Cross-references: EMBL:X65005; NID:g52643; PIDN:CAA46138.1; PID:g52644  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 24.5%; Score 26; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 7.7e-19;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASGVPARFSGSGTSLTISR 75  
|||||  
Db 52 TSNLASGVPARFSGSGTSLTISR 77

## RESULT 8

S20651  
Ig heavy chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C:Accession: S20651; S20647  
R:Losman, M.; Fasy, T.M.; Novick, K.E.; Monestier, M.  
Submitted to the EMBL Data Library, February 1992  
A:Description: Relationships among antinuclear antibodies from autoimmune MRL mice re



A:Reference number: S20639

A:Accession: S20651

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-86 <LOS>

A:Cross-references: EMBL:X65007; NID:q52647; PID:CAA46140.1; PID:q52648; EMBL:X65010; N

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 24.5%; Score 26; DB 2; Length 86;

Best Local Similarity 100.0%; Pred. No. 7.7e-19;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 TSNLASGVPARFSGSGTSLTIS 75

|||||

Db 52 TSNLASGVPARFSGSGTSLTIS 77

RESULT 9

S26339

Ig kappa chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000

C:Accession: S26339

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

J. Exp. Med. 174, 613-624, 1991

A:Title: Antibodies that are specific for a single amino acid interchange in a protein e

A:Reference number: S26309; MUID:91341421; PMID:1908510

A:Accession: S26339

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-90 <STA>

A:Cross-references: EMBL:X5195; NID:q52325; PID:CAA41905.1; PID:g1334068

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-76/Domain: immunoglobulin homology <IMM>

Query Match 24.5%; Score 26; DB 2; Length 90;

Best Local Similarity 100.0%; Pred. No. 8e-19;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 TSNLASGVPARFSGSGTSLTIS 75

|||||

Db 37 TSNLASGVPARFSGSGTSLTIS 62

RESULT 10

S17623

Ig kappa chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000

C:Accession: S17623

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17623

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-93 <CLA>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-83/Domain: immunoglobulin homology <IMM>

Query Match 24.5%; Score 26; DB 2; Length 93;

Best Local Similarity 100.0%; Pred. No. 8.2e-19;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 TSNLASGVPARFSGSGTSLTIS 75

|||||

Db 44 TSNLASGVPARFSGSGTSLTIS 69

RESULT 11

S17640

Ig kappa chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000

C:Accession: S17640

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17640

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-93 <CLA>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-83/Domain: immunoglobulin homology <IMM>

Query Match 24.5%; Score 26; DB 2; Length 93;

Best Local Similarity 100.0%; Pred. No. 8.2e-19;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 TSNLASGVPARFSGSGTSLTIS 75

|||||

Db 44 TSNLASGVPARFSGSGTSLTIS 69

RESULT 12

S17642

Ig kappa chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000

C:Accession: S17642

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17642

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-93 <CLA>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-83/Domain: immunoglobulin homology <IMM>

Query Match 24.5%; Score 26; DB 2; Length 93;

Best Local Similarity 100.0%; Pred. No. 8.2e-19;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 TSNLASGVPARFSGSGTSLTIS 75

|||||

Db 44 TSNLASGVPARFSGSGTSLTIS 69

RESULT 13

S26340

Ig light chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 19-Mar-1998 #sequence\_revision 19-Mar-1998 #text\_change 20-Jun-2000

C:Accession: S26340

R:Stark, S.E.; Caton, A.J.

J. Exp. Med. 174, 613-624, 1991

A:Title: Antibodies that are specific for a single amino acid interchange in a protei

A:Reference number: S26309; MUID:91341421; PMID:1908510

A:Accession: S26340

A:Molecule type: mRNA

A:Residues: 1-94 <STA>

A:Cross-references: EMBL:X59205; NID:q52333; PID:CAA41915.1; PID:g1334072

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:7-80/Domain: immunoglobulin homology <IMM>

Query Match 24.5%; Score 26; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 8.3e-19;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 50 TSNLACGVPARFSGSGTYSYSLTIS 75  
Db 41 TSNLACGVPARFSGSGTYSYSLTIS 66  
|||||

## RESULT 14

C33730  
Ig kappa chain V region (4.68) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Mar-1990 #sequence\_revision 18-Sep-1992 #text\_change 21-Jan-2000  
C:Accession: C33730  
R:Lawler, A.M.; Kearney, J.F.; Kuehl, M.; Gearhart, P.J.  
Proc. Natl. Acad. Sci. U.S.A. 86, 6744-6747, 1989  
A:Title: Early rearrangements of genes encoding murine immunoglobulin kappa-chains, unlinked  
A:Reference number: A33730; MUID:89367325; PMID:2505260  
A:Accession: C33730  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <LAW>  
A:Cross-references: GB:M25999; NID:gl97117; PIDN:AAA38915.1; PID:gl97118  
A>Note: the authors translated the codon TIG for residue 34 as Phe  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:16-91/Domain: immunoglobulin homology <IMM>

Query Match 24.5%; Score 26; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 8.5e-19;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 50 TSNLACGVPARFSGSGTYSYSLTIS 75  
Db 52 TSNLACGVPARFSGSGTYSYSLTIS 77  
|||||

## RESULT 15

PHI085  
Ig light chain V region (clone 163.42) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C:Accession: PHI085  
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B cell  
A:Reference number: PH0971; MUID:92381444; PMID:1512540  
A:Accession: PHI085  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-97 <TIL>  
A:Experimental source: B cell, strain (NZB x NZW)F1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:16-89/Domain: immunoglobulin homology <IMM>

Query Match 24.5%; Score 26; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 8.5e-19;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 50 TSNLACGVPARFSGSGTYSYSLTIS 75  
Db 50 TSNLACGVPARFSGSGTYSYSLTIS 75  
|||||

Search completed: November 27, 2002, 07:31:21  
Job time : 15.9614 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:23:05 ; Search time 10.2317 Seconds  
(without alignments)  
429.695 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 106

Sequence: 1 QVLSQSPAILASPGKVT.....CQWSSNPPTFGGTMLEIR 106

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	24.5	129	1	KV4A_MOUSE
2	23	21.7	107	1	KV6A_MOUSE
3	23	21.7	107	1	KV6B_MOUSE
4	23	21.7	107	1	KV6C_MOUSE
5	23	21.7	107	1	KV6D_MOUSE
6	22	20.8	107	1	KV6E_MOUSE
7	16	15.1	107	1	KV6F_MOUSE
8	16	15.1	107	1	KV6G_MOUSE
9	16	15.1	107	1	KV6H_MOUSE
10	16	15.1	107	1	KV6I_MOUSE
11	16	15.1	107	1	KV6J_MOUSE
12	14	13.2	108	1	KV3V_MOUSE
13	14	13.2	108	1	KV6K_MOUSE
14	14	13.2	111	1	KV3A_MOUSE
15	14	13.2	111	1	KV3C_MOUSE
16	14	13.2	111	1	KV3D_MOUSE
17	14	13.2	111	1	KV3E_MOUSE
18	14	13.2	111	1	KV3R_MOUSE
19	14	13.2	111	1	KV3S_MOUSE
20	14	13.2	111	1	KV3T_MOUSE
21	14	13.2	111	1	KV3U_MOUSE
22	14	13.2	112	1	KV3B_MOUSE
23	12	11.3	111	1	KV3J_MOUSE
24	12	11.3	111	1	KV3K_MOUSE
25	12	11.3	131	1	KV3I_MOUSE
26	11	10.4	109	1	KV3F_HUMAN
27	11	10.4	110	1	KV3P_MOUSE
28	11	10.4	111	1	KV3L_MOUSE
29	11	10.4	111	1	KV3M_MOUSE
30	11	10.4	111	1	KV3N_MOUSE
31	11	10.4	111	1	KV3O_MOUSE
32	11	10.4	111	1	KV3Q_MOUSE
33	11	10.4	115	1	KV3I_HUMAN

34	11	10.4	116	1	KV3J_HUMAN	P04434	homo sapien
35	11	10.4	129	1	KV3H_HUMAN	P04207	homo sapien
36	9	8.5	107	1	KV1D_HUMAN	P01596	homo sapien
37	9	8.5	108	1	KV1E_HUMAN	P01597	homo sapien
38	9	8.5	108	1	KV1H_HUMAN	P01600	homo sapien
39	9	8.5	108	1	KV1L_HUMAN	P01604	homo sapien
40	9	8.5	108	1	KV1M_HUMAN	P01605	homo sapien
41	9	8.5	108	1	KV1N_HUMAN	P01606	homo sapien
42	9	8.5	108	1	KV1O_HUMAN	P01607	homo sapien
43	9	8.5	108	1	KV1Q_HUMAN	P01609	homo sapien
44	9	8.5	108	1	KV1R_HUMAN	P01610	homo sapien
45	9	8.5	108	1	KV1S_HUMAN	P01611	homo sapien

## ALIGNMENTS

RESULT 1  
KV4A\_MOUSE STANDARD; PRT; 129 AA.  
AC P01680;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-IV region S107B precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82115300; PubMed=6799208;  
RA Kwan S.-P., Max E.E., Seidman J.G., Leder P., Scharff M.D.;  
RT "Two kappa immunoglobulin genes are expressed in the myeloma S107.";  
RL Cell 26:57-66(1981).  
CC -1- MISCELLANEOUS: THIS PROTEIN, IN WHICH THERE IS A DELETION OF TWO  
CC AMINO ACIDS AT THE V-J RECOMBINATION SITE (AFTER POSITION 118), IS  
CC SYNTHESIZED BUT NOT SECRETED IN CELLS THAT EXPRESS AND SECRETE THE  
CC NORMAL KAPPA CHAIN S107.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J00577; AAA38780.1; -;  
CC EMBL; V00780; CAA24157.1; -;  
CC PIR; A01943; KVM57B.  
CC HSSP; P01679; 2FEJ.  
CC InterPro; IPR003006; Ig\_MHC.  
CC InterPro; IPR003596; Ig\_V.  
CC Pfam; PF00047; Ig; 1.  
CC SMART; SM00406; IGV; 1.  
CC Immunoglobulin V region; Signal.  
CC SIGNAL 1 22  
CC CHAIN 23 129 IG KAPPA CHAIN V-IV REGION S107B.  
CC DOMAIN 23 45 FRAMEWORK-1.  
CC DOMAIN 46 57 COMPLEMENTARITY-DETERMINING-1.  
CC DOMAIN 58 72 FRAMEWORK-2.  
CC DOMAIN 73 79 COMPLEMENTARITY-DETERMINING-2.  
CC DOMAIN 80 111 FRAMEWORK-3.  
CC DOMAIN 112 118 COMPLEMENTARITY-DETERMINING-3.  
CC DOMAIN 119 128 FRAMEWORK-4.  
CC DISULFID 45 111 BY SIMILARITY.  
CC NON\_TER 129 129  
CC SEQUENCE 129 AA; 13833 MW; E4BB73072DCF6BE4 CRC64;  
CC -----

Query Match 24.5%; Score 26; DB 1; Length 129;  
Best Local Similarity 100.0%; Pred. No. 2.7e-19;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 50 TSNLASGVPARFSGSGTYSYSLTIS 75
      |||||
Db 74 TSNLASGVPARFSGSGTYSYSLTIS 99

RESULT 2
KV6A_MOUSE
ID KV6A_MOUSE STANDARD; PRT; 107 AA.
AC P01675;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-VI region XRPC 44.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79082830; PubMed=103573;
RA Rao D.N., Rudikoff S., Potter M.;
RT "k Chain variable regions from three galactan binding myeloma
proteins.";
RL Biochemistry 17:5555-5559(1978).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
      BIND GALACTAN.
CC PIR; A01941; KVMX4.
DR HSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
KW DOMAIN 1 23
FT DOMAIN 24 33
FT DOMAIN 34 48
FT DOMAIN 49 55
FT DOMAIN 56 87
FT DOMAIN 88 96
FT DOMAIN 97 106
FT DISULFID 23 87
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11627 MW; 27A2D022BC0A34D7 CRC64;

Query Match 21.7%; Score 23; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.6e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 LASGVPARFSGSGTYSYSLTIS 75
      |||||
Db 53 LASGVPARFSGSGTYSYSLTIS 75

RESULT 3
KV6B_MOUSE
ID KV6B_MOUSE STANDARD; PRT; 107 AA.
AC P01676;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-VI region XRPC 24.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79082830; PubMed=103573;
RA Rao D.N., Rudikoff S., Potter M.;
RT "k Chain variable regions from three galactan binding myeloma
proteins.";
RL Biochemistry 17:5555-5559(1978).

```

```

CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
      BIND GALACTAN.
CC PIR; A01941; KVMX4.
DR HSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
KW DOMAIN 1 23
FT DOMAIN 24 33
FT DOMAIN 34 48
FT DOMAIN 49 55
FT DOMAIN 56 87
FT DOMAIN 88 96
FT DOMAIN 97 106
FT DISULFID 23 87
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11584 MW; 36E6D022A5EC34D7 CRC64;

Query Match 21.7%; Score 23; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.6e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 LASGVPARFSGSGTYSYSLTIS 75
      |||||
Db 53 LASGVPARFSGSGTYSYSLTIS 75

RESULT 4
KV6C_MOUSE
ID KV6C_MOUSE STANDARD; PRT; 107 AA.
AC P01677;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-VI region TEPC 601/TEPC 191.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE (TEPC 601).
RX MEDLINE=79082830; PubMed=103573;
RA Rao D.N., Rudikoff S., Potter M.;
RT "k Chain variable regions from three galactan binding myeloma
proteins.";
RL Biochemistry 17:5555-5559(1978).
RN [2]
RP SEQUENCE (TEPC 191).
RX MEDLINE=81054757; PubMed=6776525;
RA Rudikoff S., Rao D.N., Glaudemans C.P.J., Potter M.;
RT "kappa Chain Joining segments and structural diversity of antibody
combining sites.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:4270-4274(1980).
CC -1- MISCELLANEOUS: THE TWO SEQUENCES ARE IDENTICAL.
CC -1- MISCELLANEOUS: THESE CHAINS WERE ISOLATED FROM MYELOMA PROTEINS
      THAT BIND GALACTAN.
CC PIR; A01941; KVMX4.
DR HSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
KW DOMAIN 1 23
FT DOMAIN 24 33
FT DOMAIN 34 48
FT DOMAIN 49 55
FT DOMAIN 56 87
FT DOMAIN 88 96
FT DOMAIN 97 106
FT DISULFID 23 87

```





```

SQ SEQUENCE 107 AA; 11505 MW; CA6C4284ECFCB550 CRC64;
Query Match 15.1%; Score 16; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.4e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 AEDAATYYCQWSSNP 94
KV6J_MOUSE
|||||
Db 79 AEDAATYYCQWSSNP 94

RESULT 10
KV6J_MOUSE
ID KV6J_MOUSE STANDARD; PRT; 107 AA.
AC P04943;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-VI region NQ6-8.3.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83271467; PubMed=6877353;
RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;
RT "mRNA sequences define an unusually restricted IgG response to 2-phenylloxazolone and its early diversification."
RL Nature 304:320-324(1983).
CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K00740; AAA38685.1; -
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 23
FT DOMAIN 24 33
FT DOMAIN 34 48
FT DOMAIN 49 55
FT DOMAIN 56 87
FT DOMAIN 88 96
FT DOMAIN 97 106
FT DISULFID 23 87
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11572 MW; 6F694824ECFC08E6 CRC64;

Query Match 15.1%; Score 16; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.4e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 AEDAATYYCQWSSNP 94
KV6J_MOUSE
|||||
Db 79 AEDAATYYCQWSSNP 94

RESULT 11
KV6J_MOUSE
ID KV6J_MOUSE STANDARD; PRT; 107 AA.
AC P04944;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)

```

```

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-VI region NQ5-78.2.6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83271467; PubMed=6877353;
RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;
RT "mRNA sequences define an unusually restricted IgG response to 2-phenylloxazolone and its early diversification."
RL Nature 304:320-324(1983).
CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K00744; AAA38689.1; -
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 23
FT DOMAIN 24 33
FT DOMAIN 34 48
FT DOMAIN 49 55
FT DOMAIN 56 87
FT DOMAIN 88 96
FT DOMAIN 97 106
FT DISULFID 23 87
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11611 MW; A38290781F3C30D3 CRC64;

Query Match 15.1%; Score 16; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.4e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 AEDAATYYCQWSSNP 94
KV6J_MOUSE
|||||
Db 79 AEDAATYYCQWSSNP 94

RESULT 12
KV3V_MOUSE
ID KV3V_MOUSE STANDARD; PRT; 108 AA.
AC P01674;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 2154.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin diversity."
RL Nature 276:785-790(1978).
DR PIR; A01940; KVM54.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.

```

DR Pfam: PF00047; Ig: 1.  
KW SMART; SM00406; IGV; 1.  
KW Immunoglobulin V region.  
FT DOMAIN 1 23  
FT DOMAIN 24 38  
FT DOMAIN 39 53  
FT DOMAIN 54 60  
FT DOMAIN 61 92  
FT DOMAIN 93 101  
FT DOMAIN 102 108  
FT DISULFID 23 92  
FT NON\_TER 108 108  
SQ SEQUENCE 108 AA; 11699 MW; D40921D18DAC4B9E CRC64;

Query Match 13.2%; Score 14; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.7e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 SGVPARFSGSGGT 68  
DB 60 SGVPARFSGSGGT 73

RESULT 13

ID KV6K\_MOUSE STANDARD; PRT; 108 AA.  
AC P04945;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-VI region NQ2-6.1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83271467; PubMed=6877353;  
RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
RT "mRNA sequences define an unusually restricted IgG response to 2-phenylloxazalone and its early diversification."  
RL Nature 304:320-324(1983).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; R00746; AAA38691.1; -  
DR HSSP; P01679; 2FBJ.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
KW Immunoglobulin V region; Hybridoma.  
FT DOMAIN 1 23  
FT DOMAIN 24 33  
FT DOMAIN 34 48  
FT DOMAIN 49 55  
FT DOMAIN 56 87  
FT DOMAIN 88 98  
FT DOMAIN 99 108  
FT DISULFID 23 87  
FT NON\_TER 108 108  
SQ SEQUENCE 108 AA; 11713 MW; DABF235CD9680AC6 CRC64;

Query Match 13.2%; Score 14; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.7e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 AEDAATYYCQWSS 92  
DB 79 AEDAATYYCQWSS 92

RESULT 14

ID KV3A\_MOUSE STANDARD; PRT; 111 AA.  
AC P01654;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-III region PC 2880/PC 1229.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=79073152; PubMed=103003;  
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
RT "Rearrangement of genetic information may produce immunoglobulin diversity."  
RL Nature 276:785-790(1978).  
CC -I- MISCELLANEOUS: THE PC 2880 AND PC 1229 SEQUENCES ARE IDENTICAL.  
DR PIR; A01930; KVM580.  
DR HSSP; P80362; 1WTL.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
KW Immunoglobulin V region.  
FT DOMAIN 1 23  
FT DOMAIN 24 38  
FT DOMAIN 39 53  
FT DOMAIN 54 60  
FT DOMAIN 61 92  
FT DOMAIN 93 101  
FT DOMAIN 102 111  
FT DISULFID 23 92  
FT NON\_TER 111 111  
SQ SEQUENCE 111 AA; 11980 MW; AFEAC6A9D26FC12D CRC64;

Query Match 13.2%; Score 14; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 3.8e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 SGVPARFSGSGGT 68  
DB 60 SGVPARFSGSGGT 73

RESULT 15

ID KV3C\_MOUSE STANDARD; PRT; 111 AA.  
AC P01656;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-III region MOPC 70.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=67056897; PubMed=4162931;  
RA Gray W.R., Dreyer W.J., Hood L.E.;  
RT "Mechanism of antibody synthesis: size differences between mouse kappa chains."  
RL Science 155:465-467(1967).  
CC -I- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.  
DR PIR; A01930; KVM580.



DR HSP; P01679; 2EBJ.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; Ig; 1.  
KW immunoglobulin V region; Bence-Jones protein.  
FT DOMAIN 1 23  
FT DOMAIN 24 38  
FT DOMAIN 39 53  
FT DOMAIN 54 60  
FT DOMAIN 61 92  
FT DOMAIN 93 101  
FT DOMAIN 102 111  
FT DISULFID 23 92  
FT NON\_TER 111 111  
SQ SEQUENCE 111 AA; 11904 MW; 4FE7ABC9DF0FC125 CRC64;

Query Match 13.2%; Score 14; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 3.8e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 SGVPARFSGSGST 68  
Db 60 SGVPARFSGSGST 73

Search completed: November 27, 2002, 07:28:47  
Job time : 10.2317 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: November 27, 2002, 07:24:25 : Search time: 29.8764 Seconds  
(without alignments)  
731.044 Million cell updates/sec

Title: US-09-893-615-89  
Perfect score: 106  
Sequence: 1 QIVLSQSPAILASPGKVT.....CQWSSNPPTGGGTMLEIR 106

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 671580 seqs, 206047115 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	41.5	97	Q9JL76	Q9JL76 mus musculus
2	23	21.7	134	Q8VDD0	Q8VDD0 mus musculus
3	17	16.0	106	Q9U410	Q9U410 schistosoma
4	16	15.1	235	Q9JL12	Q9JL12 mus musculus
5	15	14.2	101	Q9JL78	Q9JL78 mus musculus
6	14	13.2	103	Q9JL80	Q9JL80 mus musculus
7	14	13.2	111	Q920E9	Q920E9 mus musculus
8	11	10.4	108	Q9UL83	Q9UL83 homo sapien
9	11	10.4	109	Q9UL85	Q9UL85 homo sapien
10	9	8.5	104	Q9JL82	Q9JL82 mus musculus
11	9	8.5	107	Q9UL81	Q9UL81 homo sapien
12	9	8.5	107	Q96SA9	Q96SA9 homo sapien
13	9	8.5	108	Q9UL79	Q9UL79 homo sapien
14	9	8.5	108	Q9UL77	Q9UL77 homo sapien
15	9	8.5	108	Q9UL70	Q9UL70 homo sapien
16	9	8.5	109	Q9UL78	Q9UL78 homo sapien

17	9	8.5	109	6	Q9NOW5	Q9NOW5 oryctolagus
18	9	8.5	109	11	Q9ET13	Q9ET13 mus musculus
19	9	8.5	109	11	Q920E6	Q920E6 mus musculus
20	9	8.5	114	4	Q9UL80	Q9UL80 homo sapien
21	9	8.5	129	11	Q8VDE2	Q8VDE2 mus musculus
22	9	8.5	233	11	Q9IWS9	Q9IWS9 mus musculus
23	9	8.5	234	11	Q9IWF8	Q9IWF8 mus musculus
24	9	8.5	234	11	Q8R062	Q8R062 mus musculus
25	9	8.5	238	11	Q9M37	Q9M37 mus musculus
26	9	8.5	238	11	Q8VC16	Q8VC16 mus musculus
27	9	8.5	239	4	Q8TCD0	Q8TCD0 homo sapien
28	9	8.5	239	11	Q8VC55	Q8VC55 mus musculus
29	9	8.5	298	11	Q9QYF0	Q9QYF0 mus musculus
30	8	7.5	107	11	Q9JL84	Q9JL84 mus musculus
31	8	7.5	214	11	Q9JL85	Q9JL85 mus musculus
32	8	7.5	241	11	Q921A6	Q921A6 mus musculus
33	8	7.5	338	5	Q8T8J5	Q8T8J5 drosophila
34	8	7.5	385	5	Q9NHV3	Q9NHV3 drosophila
35	8	7.5	385	5	Q9NHV2	Q9NHV2 drosophila
36	8	7.5	385	5	Q9NHV1	Q9NHV1 drosophila
37	8	7.5	385	5	Q9NHV0	Q9NHV0 drosophila
38	8	7.5	385	5	Q9NHU9	Q9NHU9 drosophila
39	8	7.5	385	5	Q9NHU8	Q9NHU8 drosophila
40	8	7.5	385	5	Q9NHU7	Q9NHU7 drosophila
41	8	7.5	385	5	Q9NHU6	Q9NHU6 drosophila
42	8	7.5	385	5	Q9NHU5	Q9NHU5 drosophila
43	8	7.5	385	5	Q9N666	Q9N666 drosophila
44	8	7.5	385	5	Q8T8L2	Q8T8L2 drosophila
45	8	7.5	385	5	Q8T8L0	Q8T8L0 drosophila

## ALIGNMENTS

RESULT 1

Q9JL76	PRELIMINARY;	PRT;	97 AA.
ID	Q9JL76		
AC	Q9JL76;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	Anti-mycosin immunoglobulin light chain variable region (Fragment).		
DE	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=DBA/2;		
RX	MEDLINE=20448942; PubMed=10992488;		
RA	Walkiel S., Liao L., Cunningham M.W., Diamond B.:		
RT	T-cell-dependent antibody response to the dominant epitope of streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive with cardiac myosin.		
RL	Infect. Immun. 68:5803-5808(2000).		
DR	EMBL; AF206030; AAF69328.1;		
DR	HSSP; P01679; 2FBJ		
DR	InterPro; IPR003006; Ig_MHC.		
DR	InterPro; IPR003596; Ig_V.		
DR	Pfam; PF00047; Ig_1.		
DR	SMART; SM00406; IGV; 1.		
FT	NON_TER 1		
FT	NON_TER 97 97		
SQ	SEQUENCE 97 AA; 10542 MW; C9EE1FFELF49DALC CKC64;		

Query Match 41.5%; Score 44; DB 11; Length 97;

Best Local Similarity 100.0%; Pred. No. 1.5e-37;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 49 ATSNLASGVPARTSGSGTGYSLTISRVEADAATYTCQWSS 92

Db 39 ATSNLASGVPARTSGSGTGYSLTISRVEADAATYTCQWSS 82

```

RESULT 2
Q8VDD0      PRELIMINARY;      PRT;      134 AA.
AC Q8VDD0;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Anti-MOG 212 variable light chain (Fragment).
GN ANTI-MOG KAPPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Chernajovsky Y.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Sembli P.;
RT "targeting T cells to the CNS.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ416331; CAC94866.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 14525 MW; CFDF8E2236E2D0CF CRC64;

Query Match      21.7%; Score 23; DB 11; Length 134;
Best Local Similarity 100.0%; Pred. No. 9.4e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 LASGVPARFSGSGTSYSLTIS 75
   |||||
DB 75 LASGVPARFSGSGTSYSLTIS 97

RESULT 3
Q9U410      PRELIMINARY;      PRT;      106 AA.
AC Q9U410;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Monoclonal anti-idiotypic antibody NP30 immunoglobulin light chain
DE variable region (Fragment).
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RA Song X.T.; Feng Z.Q.; Qiu Z.N.; Li Y.Q.; Huang H.L.; Guan X.H.;
RT "Amplification, cloning and sequence analysis of the light chain
RT variable region gene of monoclonal anti-idiotypic antibody NP30 of
RT Schistosoma japonicum.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207620; AAF19434.1; -.
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1 1
FT NON_TER 106 106

```

```

SQ SEQUENCE 106 AA; 11478 MW; F20F544426BAE63E CRC64;

Query Match      16.0%; Score 17; DB 5; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RFGSGSGTSYSLTISR 76
   |||||
DB 60 RFGSGSGTSYSLTISR 76

RESULT 4
Q9IWI2      PRELIMINARY;      PRT;      235 AA.
AC Q9IWI2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:6582).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST TUMOR;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006643; AAH06643.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 235 AA; 26021 MW; 5FC73BDEBD5E8FEF CRC64;

Query Match      15.1%; Score 16; DB 11; Length 235;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RFGSGSGTSYSLTIS 75
   |||||
DB 82 RFGSGSGTSYSLTIS 97

RESULT 5
Q9JL78      PRELIMINARY;      PRT;      101 AA.
AC Q9JL78;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Anti-myosin immunoglobulin light chain variable region
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A.CA;
RX MEDLINE=20448942; PubMed=10992488;
RA Maiktel S.; Liao L.; Cunningham M.W.; Diamond B.;
RT "T-Cell-dependent antibody response to the dominant epitope of
RT streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive
RT with cardiac myosin.";
RL Infect. Immun. 68:5803-5808 (2000).
DR EMBL; AF206028; AAF69326.1; -.
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1 1
FT NON_TER 101 101

```



RT fetus.;  
RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
DR EMBL; AF035029; AAD56265.1; -.  
DR HSSP; P80362; 1WTL.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
FT NON\_TER 1  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 11761 MW; FB1E43E7C7AFACCC CRC64;

Query Match 10.4%; Score 11; DB 4; Length 109;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 PARFSGSGSGT 68  
Db 59 PARFSGSGSGT 69

## RESULT 10

Q9JL82 Q9JL82 PRELIMINARY; PRT; 104 AA.  
AC Q9JL82;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Anti-myosin immunoglobulin light chain variable region  
DE (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C;  
RX MEDLINE=20448942; PubMed=1092488;  
RA Malkiel S., Liao L., Cunningham M.W., Diamond B.;  
RT "T-Cell-dependent antibody response to the dominant epitope of  
streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive  
with cardiac myosin";  
RL Infect. Immun. 68:5803-5808(2000).  
DR EMBL; AF206024; AAF69322.1; -.  
DR HSSP; P01607; IREI.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
FT NON\_TER 1  
FT NON\_TER 104  
SQ SEQUENCE 104 AA; 11360 MW; 5DA8BBFD5FOAJAE CRC64;

Query Match 8.5%; Score 9; DB 11; Length 104;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 RFGSGSGSGT 68  
Db 58 RFGSGSGSGT 66

## RESULT 11

Q9JL81 Q9JL81 PRELIMINARY; PRT; 107 AA.  
AC Q9JL81;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Myosin-reactive immunoglobulin light chain variable region  
DE (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98277139; PubMed=9614934;  
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,  
RA Young D.C.;  
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
fetus";  
RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
DR EMBL; AF035033; AAD56269.1; -.  
DR HSSP; P01607; IREI.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
FT NON\_TER 1  
FT NON\_TER 107  
SQ SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;

Query Match 8.5%; Score 9; DB 4; Length 107;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 RFGSGSGSGT 68  
Db 61 RFGSGSGSGT 69

## RESULT 12

Q96SA9 Q96SA9 PRELIMINARY; PRT; 107 AA.  
AC Q96SA9;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Anti-streptococcal/anti-myosin immunoglobulin kappa light chain  
DE variable region (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98375893; PubMed=9712075;  
RA Adderson E.E., Shikhan A.R., Ward K.E., Cunningham M.W.;  
RT "Molecular analysis of polyclonal antibodies from  
rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin  
antibody V region genes";  
RL J. Immunol. 161:2020-2031(1998).  
DR EMBL; U96396; AAB68785.1; -.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 1.  
FT NON\_TER 1  
FT NON\_TER 107  
SQ SEQUENCE 107 AA; 11520 MW; 4BB43E9C5B577F16 CRC64;

Query Match 8.5%; Score 9; DB 4; Length 107;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 RFGSGSGSGT 68  
Db 61 RFGSGSGSGT 69

## RESULT 13

Q9JL79 Q9JL79 PRELIMINARY; PRT; 108 AA.  
AC Q9JL79;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035035; AAD56273.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT SEQUENCE 108 AA; 11787 MW; DB5845F19724FB4E CRC64;
SQ
Query Match 8.5%; Score 9; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 RFSGSGSGT 68
DB 61 RFSGSGSGT 69
RESULT 14
Q9UL77 PRELIMINARY; PRT; 108 AA.
AC Q9UL77;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035037; AAD56273.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT SEQUENCE 108 AA; 11738 MW; C06681716C4D16F3 CRC64;
SQ
Query Match 8.5%; Score 9; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 RFSGSGSGT 68
DB 61 RFSGSGSGT 69
```

```
RESULT 15
Q9UL70 PRELIMINARY; PRT; 108 AA.
AC Q9UL70;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035044; AAD56280.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT SEQUENCE 108 AA; 11633 MW; B7BEDC3E41FCCA37 CRC64;
SQ
Query Match 8.5%; Score 9; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 RFSGSGSGT 68
DB 61 RFSGSGSGT 69
Search completed: November 27, 2002, 07:30:27
Job time : 31.1264 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:26:06 ; Search time 14.7336 Seconds  
(without alignments)  
211.682 Million cell updates/sec

Title: US-09-893-615-89

Perfect score: 106

Sequence: 1 QIVLSQSPAILASPGKVT.....CQWSSNPPTFGGCTMLEIR 106

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/PTIUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	44.3	106	3	US-08-783-853A-105
2	47	44.3	106	4	US-09-344-050-105
3	47	44.3	107	3	US-08-783-853A-11
4	47	44.3	107	4	US-09-344-050-11
5	47	44.3	112	3	US-08-783-853A-103
6	47	44.3	112	4	US-09-344-050-103
7	40	37.7	129	2	US-08-449-287-2
8	40	37.7	235	4	US-09-423-439-18
9	40	37.7	235	4	US-09-423-439-58
10	40	37.7	235	4	US-09-011-769A-23
11	31	29.2	128	1	US-08-476-275-4
12	31	29.2	128	4	US-08-475-815B-7
13	27	25.5	105	3	US-08-434-000A-12
14	27	25.5	105	4	US-09-312-157-12
15	27	25.5	106	1	US-07-634-278-1
16	27	25.5	106	1	US-07-634-278-16
17	27	25.5	106	1	US-08-477-728-1
18	27	25.5	106	1	US-08-477-728-16
19	27	25.5	106	1	US-08-474-040-1
20	27	25.5	106	1	US-08-474-040-16
21	27	25.5	106	1	US-08-487-200-1
22	27	25.5	106	1	US-08-487-200-16
23	27	25.5	106	1	US-08-488-113B-163
24	27	25.5	106	1	US-08-477-484B-163
25	27	25.5	106	1	US-08-107-669D-49
26	27	25.5	106	1	US-08-472-788A-83
27	27	25.5	106	2	US-08-477-531B-49

28	27	25.5	106	2	US-08-646-360-163	Sequence 163, Appl
29	27	25.5	106	2	US-08-082-842A-83	Sequence 83, Appl
30	27	25.5	106	4	US-08-839-765-163	Sequence 163, Appl
31	27	25.5	106	4	US-09-136-389-163	Sequence 163, Appl
32	27	25.5	106	4	US-08-484-537-1	Sequence 1, Appl
33	27	25.5	106	4	US-08-484-537-16	Sequence 16, Appl
34	27	25.5	106	4	US-09-610-838-163	Sequence 163, Appl
35	27	25.5	107	1	US-08-211-202-3	Sequence 3, Appl
36	27	25.5	108	4	US-09-171-945-9	Sequence 9, Appl
37	27	25.5	129	2	US-08-116-778E-2	Sequence 2, Appl
38	27	25.5	129	2	US-08-438-562-2	Sequence 2, Appl
39	27	25.5	129	2	US-08-483-528B-92	Sequence 92, Appl
40	27	25.5	130	4	US-09-393-385B-113	Sequence 113, Appl
41	27	25.5	230	4	US-09-485-737B-102	Sequence 102, Appl
42	27	25.5	233	4	US-09-485-737B-69	Sequence 69, Appl
43	27	25.5	235	4	US-09-171-945-17	Sequence 17, Appl
44	27	25.5	235	4	US-09-485-737B-93	Sequence 93, Appl
45	27	25.5	239	3	US-08-279-772A-8	Sequence 8, Appl

#### ALIGNMENTS

RESULT 1  
US-08-783-853A-105  
; Sequence 105, Application US/08783853A  
; Patent No. 6005091  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; APPLICANT: Church, William  
; APPLICANT: Gross, Mitchell  
; APPLICANT: Feuerstein, Giora  
; APPLICANT: Nichols, Andrew  
; APPLICANT: Padlan, Eduardo  
; APPLICANT: Patel, Arunbhai  
; APPLICANT: Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; NUMBER OF SEQUENCES: 111  
; TITLE OF INVENTION: OF THROMBOSIS  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/783,853A  
; FILING DATE: 16-JAN-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/029,119  
; FILING DATE: 24-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 105:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

```

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-783-853A-105

```

Query Match 44.3%; Score 47; DB 3; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.8e-34;  
Matches 47; Conservative 0; Mismatches 0; Indels

Qy 1 QIVLSQSPAILSASPGKEKVTMTCRASSSVNYMHWYQQKPGSSPKPWI 47  
|||||  
Db 1 QIVLSQSPAILSASPGKEKVTMTCRASSSVNYMHWYQQKPGSSPKPWI 47  
|||||

## RESULT 2

US-09-344-050-105  
; Sequence 105, Application US/09344050  
; Patent No. 6391299

/ Applicant No: 0031933  
 / GENERAL INFORMATION:  
 / APPLICANT: Blackburn, Michael  
 / APPLICANT: Church, William  
 / APPLICANT: Gross, Mitchell  
 / APPLICANT: Feuerstein, Giora  
 / APPLICANT: Nichols, Andrew  
 / APPLICANT: Padlan, Eduardo  
 / APPLICANT: Patel, Arunbhai  
 / APPLICANT: Sylvester, Daniel

APPLICANT: SYLVESCU, DANIEL  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS  
TITLE OF INVENTION: OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:

ADDRESS: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia

```

: STATE: PA
: COUNTRY: USA
: ZIP: 19406
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSEQ Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/344,050
: FILING DATE: 24-JUN-1999

```

CLASSIFICATION: 1  
PRIOR APPLICATION DATA: 2  
APPLICATION NUMBER: 08/783,853 3  
FILING DATE: 16-JAN-1997 4  
ATTORNEY/AGENT INFORMATION: 5  
NAME: Baunelster, Kirk 6  
REGISTRATION NUMBER: 33,833 7  
REFERENCE/DOCKET NUMBER: P50438 8  
TELECOMMUNICATION INFORMATION: 9  
TELEPHONE: 610-270-5096 10

```

; ELEX:
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids

```

TYPE: amil;  
STRANDEDNESS:  
TOPOLOGY:  
MOLECULE TYPE  
HYPOTHETICAL  
ANTI-SENSE:  
FRAGMENT TYPE  
ORIGINAL SOURCE:  
US-09-344-050-105

Matches

Qy

Db

### RESULT 3

US-08-783-853A-11  
; Sequence 11, Application US/08783853A  
; Patent No. 6005091

;  
 ; GENERAL INFORMATION:  
 ;  
 ; APPLICANT: Blackburn, Michael  
 ; APPLICANT: Church, William  
 ; APPLICANT: Gross, Mitchell  
 ; APPLICANT: Feuerstein, Giora  
 ; APPLICANT: Nichols, Andrew  
 ; APPLICANT: Padlan, Eduardo  
 ; APPLICANT: Patel, Arunbhai  
 ;

```

; REFERENCE: Fater, Abdulmalik
; APPLICANT: Sylvester, Daniel
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT
; OF THROMBOSIS
; TITLE OF INVENTION: OF THROMBOSIS
; NUMBER OF SEQUENCES: 111

```

NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESS: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road

STREET: 709 S. Swedeland Road  
 CITY: King of Prussia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19406  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/783,853A  
 FILING DATE: 16-JAN-1997

CLASSIFICATION: 10 JAN 1997  
PRIOR APPLICATION DATA: 60/029,119  
APPLICATION NUMBER: 60/029,119  
FILING DATE: 24-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:

```

?
? INFORMATION FOR
? SEQUENCE CHA
? LENGTH: 1
? TYPE: ami
? STRANDEDNE
? TOPOLOGY:
? MOLECULE TYPE
? HYPOTHETICAL
? ANTI-SENSE:
? FRAGMENT TYPE
? ORIGINAL SOURCE
? US-08-783-953A-11

```

Qy Db

## RESULT 4

US-09-344-050-11  
; Sequence 11, Application US/09344050  
; Patent No. 6391299  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; APPLICANT: Church, William  
; APPLICANT: Gross, Mitchell  
; APPLICANT: Feuerstein, Giora  
; APPLICANT: Nichols, Andrew  
; APPLICANT: Padlan, Eduardo  
; APPLICANT: Patel, Arunbhai  
; APPLICANT: Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/344,050  
; FILING DATE: 24-JUN-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/783,853  
; FILING DATE: 16-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 107 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-09-344-050-11

Query Match 44.3%; Score 47; DB 4; Length 107;  
Best Local Similarity 100.0%; Pred. No. 1.9e-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSQSPAILSPGSEKVTMTCRASSSYNMYHWYQKPGSSPKPWI 47  
|||||  
DB 1 QIVLSQSPAILSPGSEKVTMTCRASSSYNMYHWYQKPGSSPKPWI 47

RESULT 5  
US-08-783-853A-103  
; Sequence 103, Application US/08783853A  
; Patent No. 6005091  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; APPLICANT: Church, William  
; APPLICANT: Gross, Mitchell  
; APPLICANT: Feuerstein, Giora

; APPLICANT: Nichols, Andrew  
; APPLICANT: Padlan, Eduardo  
; APPLICANT: Patel, Arunbhai  
; APPLICANT: Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/783,853A  
; FILING DATE: 16-JAN-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/029,119  
; FILING DATE: 24-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: P50438  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 103:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 112 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-08-783-853A-103

Query Match 44.3%; Score 47; DB 3; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.9e-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSQSPAILSPGSEKVTMTCRASSSYNMYHWYQKPGSSPKPWI 47  
|||||  
DB 1 QIVLSQSPAILSPGSEKVTMTCRASSSYNMYHWYQKPGSSPKPWI 47

RESULT 6  
US-09-344-050-103  
; Sequence 103, Application US/09344050  
; Patent No. 6391299  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; APPLICANT: Church, William  
; APPLICANT: Gross, Mitchell  
; APPLICANT: Feuerstein, Giora  
; APPLICANT: Nichols, Andrew  
; APPLICANT: Padlan, Eduardo  
; APPLICANT: Patel, Arunbhai  
; APPLICANT: Sylvester, Daniel  
; TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,050  
FILING DATE: 24-JUN-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-09-344-050-103

Query Match 44.3%; Score 47; DB 4; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.9e-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QIVLSQSPAILASPGKVTMTCRASSVNMHWYQKPGSSPKPWI 47  
|||||  
Db 1 QIVLSQSPAILASPGKVTMTCRASSVNMHWYQKPGSSPKPWI 47

RESULT 7  
US-08-449-287-2  
Sequence 2, Application US/08449287  
Patent No. 5877293  
GENERAL INFORMATION:  
APPLICANT: ADAIR, John Robert  
APPLICANT: BOWMER, Mark William  
APPLICANT: MOUNTAIN, Andrew  
APPLICANT: OWENS, Raymond John  
TITLE OF INVENTION: CDR Grafted Anti-CEA Antibodies and  
TITLE OF INVENTION: Their Production  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,287

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/154,389  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT GB91/01108  
FILING DATE: 05-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9014932.9  
FILING DATE: 05-JUL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT GB90/02017  
FILING DATE: 21-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: SAXE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 40283/110 CARA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 129 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-449-287-2

Query Match 37.7%; Score 40; DB 2; Length 129;  
Best Local Similarity 100.0%; Pred. No. 3.2e-28;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 ATSNLASGVPARFSGSGTSTSLTISRVEADAATYYCQ 88  
|||||  
Db 71 ATSNLASGVPARFSGSGTSTSLTISRVEADAATYYCQ 110

RESULT 8  
US-09-423-439-18  
Sequence 18, Application US/09423439  
Patent No. 6339070  
GENERAL INFORMATION:  
APPLICANT: EMERY, Stephen Charles  
APPLICANT: BLAKEY, David Charles  
TITLE OF INVENTION: CHEMICAL COMPOUNDS  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Winthrop, L.L.P.  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/423,439  
FILING DATE: 09-No. 6339070-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB98/01294  
FILING DATE: 05-MAY-1998  
APPLICATION NUMBER: GB 9709421.3  
FILING DATE: 10-MAY-1997  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 235 amino acids  
TYPE: amino acid

```
;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-423-439-18

Query Match          37.7%; Score 40; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 5.3e-28;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQ 88
Db 71 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQ 110

RESULT 9
US-09-423-439-58
; Sequence 58, Application US/09423439
; Patent No. 6339070
; GENERAL INFORMATION:
; APPLICANT: EMERY, Stephen Charles
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Winthrop, L.L.P.
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01294
; FILING DATE: 05-MAY-1998
; APPLICATION NUMBER: GB 9709421.3
; FILING DATE: 10-MAY-1997
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-423-439-58

Query Match          37.7%; Score 40; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 5.3e-28;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQ 88
Db 71 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQ 110

RESULT 10
US-09-011-769A-23
; Sequence 23, Application US/09011769A
; Patent No. 6436691
; GENERAL INFORMATION:
; APPLICANT: SLATER, Anthony M.
; BLAKEY, David C.
; DAVIES, David H.
; HENNAW, John F.
; HENNEQUIN, Laurent F.A.
; MARSHAM, Peter R.
; DOWELL, Robert I.
; TITLE OF INVENTION: Chemical Compounds
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, LLP
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011.769A
; FILING DATE: 13-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/01975
; FILING DATE: 13-AUG-1996
; APPLICATION NUMBER: GB 9612295.7
; FILING DATE: 12-JUN-1996
; APPLICATION NUMBER: GB 9611019.2
; FILING DATE: 25-MAY-1996
; APPLICATION NUMBER: GB 9516810.0
; FILING DATE: 16-AUG-1995
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-011-769A-23

Query Match          37.7%; Score 40; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 5.3e-28;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQ 88
Db 71 ATSNLASGVPARFSGSGTSLTISRVEADAATYYCQ 110

RESULT 11
US-08-476-275-4
; Sequence 4, Application US/08476275
; Patent No. 5776456
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; APPLICANT: Hanna, Nabil
; APPLICANT: Leonard, John E.
; APPLICANT: Newman, Roland A.
; APPLICANT: Reff, Mitchell E.
; APPLICANT: Rastetter, William H.
; TITLE OF INVENTION: Therapeutic Application of Chimeric and
; Radiolabeled Antibodies to Human B Lymphocyte Restricted
; Differentiation Antigen for the Treatment of B-Cell
; Lymphoma
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince St.
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,275  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
APPLICATION NUMBER: US 07/978,891  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-155  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 128 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-275-4

Query Match 29.2%; Score 31; DB 1; Length 128;  
Best Local Similarity 100.0%; Pred. No. 2.7e-20;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RFGSGSGTSYSLTISRVEADAATYYCQW 90  
|||||  
DB 82 RFGSGSGTSYSLTISRVEADAATYYCQW 112

## RESULT 12

US-08-475-815B-7  
Sequence 7, Application US/08475815B  
Patent No. 6399061  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DARRELL R.  
APPLICANT: HANNA, NABIL  
APPLICANT: LEONARD, JOHN E.  
APPLICANT: NEWMAN, ROLAND A.  
APPLICANT: REFF, MITCHELL E.  
APPLICANT: RASTETER, WILLIAM H.  
TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND  
TITLE OF INVENTION: RADIOLABELLED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED  
TITLE OF INVENTION: DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL  
TITLE OF INVENTION: LYMPHOMA  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY WINTHROP  
STREET: 1100 New York Avenue, N.W., Ninth FL.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,815B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/978,891

FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 23522-0157  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 128 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-475-815B-7  
Query Match 29.2%; Score 31; DB 4; Length 128;  
Best Local Similarity 100.0%; Pred. No. 2.7e-20;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RFGSGSGTSYSLTISRVEADAATYYCQW 90  
|||||  
DB 82 RFGSGSGTSYSLTISRVEADAATYYCQW 112

## RESULT 13

US-08-434-000A-12  
Sequence 12, Application US/08434000A  
Patent No. 6046037  
GENERAL INFORMATION:  
APPLICANT: ANDREW C. HIATT, JULIAN  
APPLICANT: K.-C. MA, THOMAS LEHNER  
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,000A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below: 1  
APPLICATION NUMBER: 08/367,395  
FILING DATE: 12/30/94  
ATTORNEY/AGENT INFORMATION:  
NAME: Guise, Jeffrey W.  
REGISTRATION NUMBER: 34,613  
REFERENCE/DOCKET NUMBER: 212/127  
TELEPHONE: (619) 552-8400  
TELEFAX: (619) 552-0159  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

TOPOLOGY: DESCRIPTION: Guy's 13 Kappa  
US-08-434-000A-12

Query Match 25.5%; Score 27; DB 3; Length 105;  
Best Local Similarity 100.0%; Pred. No. 7.5e-17;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASCVPARFSGSGTYSLTISR 76  
|||||

DB 50 TSNLASCVPARFSGSGTYSLTISR 76  
|||||

## RESULT 14

US-09-312-157-12  
Sequence 12, Application US/09312157

Patent No. 6303341

## GENERAL INFORMATION:

APPLICANT: ANDREW C. HIATT, JULIAN  
K.-C. MA, THOMAS LEHNER

TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
PROTEINS IN PLANTS AND THEIR USES

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/312,157

FILING DATE: 14-May-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/434,000

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Guise, Jeffrey W.

REGISTRATION NUMBER: 34,613

REFERENCE/DOCKET NUMBER: 212/127

TELEPHONE: (619) 552-8400

TELEFAX: (619) 552-0159

TELEX: 67-351

SEQUENCE LISTING

INFORMATION FOR SEQ ID NO: 12

SEQUENCE CHARACTERISTICS:

LENGTH: 105 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

DESCRIPTION: Guy's 13 Kappa

US-09-312-157-12

Query Match 25.5%; Score 27; DB 4; Length 105;

Best Local Similarity 100.0%; Pred. No. 7.5e-17;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASCVPARFSGSGTYSLTISR 76  
|||||

DB 50 TSNLASCVPARFSGSGTYSLTISR 76  
|||||

## RESULT 15

US-07-634-278-1

Sequence 1, Application US/07634278  
Patent No. 5530101

## GENERAL INFORMATION:

APPLICANT: QUEEN, Cary L.

APPLICANT: CO, Man Sung

APPLICANT: SCHNEIDER, William P.

APPLICANT: LANDOLFI, Nicholas F.

APPLICANT: COELINGH, Kathleen L.

APPLICANT: SELICK, Harold E.

TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: 379 Lytton Avenue

CITY: Palo Alto

STATE: California

COUNTRY: US

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/634,278

FILING DATE: 19-DEC-1990

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/590,274

FILING DATE: 28-SEP-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/310,252

FILING DATE: 13-FEB-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/290,975

FILING DATE: 28-DEC-1988

ATTORNEY/AGENT INFORMATION:

NAME: Smith, William M

REGISTRATION NUMBER: 30,223

REFERENCE/DOCKET NUMBER: 11823-002600

TELEPHONE: (415) 326-2400

TELEFAX: (415) 326-2422

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: Protein

LOCATION: 1..106

OTHER INFORMATION: /note= "Variable region of the mouse

OTHER INFORMATION: anti-tac antibody light chain."

US-07-634-278-1

Query Match 25.5%; Score 27; DB 1; Length 106;

Best Local Similarity 100.0%; Pred. No. 7.5e-17;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASCVPARFSGSGTYSLTISR 76  
|||||

DB 50 TSNLASCVPARFSGSGTYSLTISR 76  
|||||

Search completed: November 27, 2002, 07:32:11

Job time : 15.7336 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:28:06 ; Search time 9.82239 Seconds  
(without alignments)  
171.849 Million cell updates/sec

Title: US-09-893-615-89  
Perfect score: 106  
Sequence: 1 QIVLSQSPAILASPGKVT.....CQWSSNPPTFGGTMLEIR 106

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 102317 seqs, 15924203 residues

Word size : 0

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	106	10	US-09-893-615-89
2	47	44.3	106	10	US-09-893-615-89
3	47	44.3	106	12	US-10-051-852-105
4	47	44.3	107	10	US-09-965-099-11
5	47	44.3	107	12	US-10-051-852-11
6	47	44.3	112	10	US-09-965-099-103
7	47	44.3	112	12	US-10-051-852-103
8	27	25.5	105	9	US-09-982-107-12
9	27	25.5	106	10	US-09-976-787-24
10	27	25.5	106	10	US-09-865-198-23
11	27	25.5	107	9	US-09-144-886-88
12	27	25.5	108	10	US-09-976-787-8
13	27	25.5	108	10	US-09-865-198-8
14	27	25.5	108	10	US-09-910-059-9
15	27	25.5	112	9	US-09-144-886-89
16	27	25.5	235	10	US-09-910-059-17
17	27	25.5	238	10	US-09-976-787-29
18	27	25.5	238	10	US-09-865-198-28
19	27	25.5	240	10	US-09-976-787-28

Sequence 27, Appl  
Sequence 2, Appl  
Sequence 76, Appl  
Sequence 92, Appl  
Sequence 98, Appl  
Sequence 82, Appl  
Sequence 83, Appl  
Sequence 28, Appl  
Sequence 6, Appl  
Sequence 57, Appl  
Sequence 62, Appl  
Sequence 57, Appl  
Sequence 62, Appl  
Sequence 78, Appl  
Sequence 78, Appl  
Sequence 66, Appl  
Sequence 66, Appl  
Sequence 66, Appl  
Sequence 74, Appl  
Sequence 95, Appl  
Sequence 95, Appl  
Sequence 99, Appl  
Sequence 99, Appl  
Sequence 91, Appl  
Sequence 80, Appl

20 27 25.5 240 10 US-09-865-198-27  
21 27 25.5 669 9 US-09-807-721-2  
22 26 24.5 107 9 US-09-144-886-76  
23 25 23.6 109 9 US-09-144-886-92  
24 25 23.6 109 9 US-09-144-886-98  
25 23 21.7 107 9 US-09-144-886-82  
26 23 21.7 107 9 US-09-144-886-83  
27 23 21.7 119 10 US-09-808-037-28  
28 23 21.7 239 10 US-09-808-037-6  
29 20 18.9 107 10 US-09-965-099-57  
30 20 18.9 107 10 US-09-965-099-62  
31 20 18.9 107 12 US-10-051-852-57  
32 20 18.9 107 12 US-10-051-852-62  
33 20 18.9 125 10 US-09-965-099-78  
34 20 18.9 125 12 US-10-051-852-78  
35 19 17.9 93 10 US-09-965-099-66  
36 19 17.9 93 12 US-10-051-852-66  
37 19 17.9 107 10 US-09-965-099-74  
38 19 17.9 107 12 US-10-051-852-74  
39 19 17.9 109 10 US-09-965-099-95  
40 19 17.9 109 12 US-10-051-852-95  
41 19 17.9 129 10 US-09-965-099-99  
42 19 17.9 129 12 US-10-051-852-99  
43 18 17.0 131 10 US-09-881-823-6  
44 18 17.0 109 9 US-09-144-886-91  
45 18 17.0 127 10 US-09-753-436-80

## ALIGNMENTS

RESULT 1  
US-09-893-615-89  
; Sequence 89, Application US/09893615  
; Patent No. US20020082395A1  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Gerald W.  
; Wong, Hing  
; Stinson, Jeffrey L.  
; TITLE OF INVENTION: CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
; POSITIVE BACTERIA  
; NUMBER OF SEQUENCES: 89  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
; DUNNER, LLP  
; STREET: 1300 I Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/893,615  
; FILING DATE: 29-Jun-2001  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Einaudi, Carol P.  
; REGISTRATION NUMBER: 32,220  
; REFERENCE/DOCKET NUMBER: 04995.0041-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 89:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 89:  
US-09-893-615-89

Query Match 100.0%; Score 106; DB 10; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.8e-86;  
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPWISATSNLASGVPAR 60  
DB 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPWISATSNLASGVPAR 60

QY 61 FSGSGSGTSYLSIRVEADAATYYCQWSSNPPFTFGGTMLEIR 106  
DB 61 FSGSGSGTSYLSIRVEADAATYYCQWSSNPPFTFGGTMLEIR 106

RESULT 2

US-09-965-099-105  
; Sequence 105, Application US/09965099  
; Patent No. US20020136725A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Feuerstein, Giora  
Patel, Arunbhai

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 105:

SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 105:

US-09-965-099-105

Query Match 44.3%; Score 47; DB 10; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1e-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47

DB 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47

RESULT 3

US-10-051-852-105  
; Sequence 105, Application US/10051852  
; Patent No. US20020146411A1

GENERAL INFORMATION:

APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Giora  
Nichols, Andrew  
Padlan, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/051,852  
FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 105:

SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 105:

US-10-051-852-105

Query Match 44.3%; Score 47; DB 12; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1e-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47

DB 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47

RESULT 4

US-09-965-099-11  
; Sequence 11, Application US/09965099  
; Patent No. US20020136725A1  
; GENERAL INFORMATION:  
; APPLICANT: Blackburn, Michael  
; Feuerstein, Giora  
; Patel, Arunbhai

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN  
TREATMENT OF THROMBOSIS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/965,099

FILING DATE: 26-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/346,487

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: P50438-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096

TELEFAX: <Unknown>

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-965-099-11  
Query Match 44.3%; Score 47; DB 10; Length 107;  
Best Local Similarity 100.0%; Pred. No. le-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QIVLSQSPAILSPGKVTMTCRASSVNYMHVYQKPGSSPKPWI 47

Db 1 QIVLSQSPAILSPGKVTMTCRASSVNYMHVYQKPGSSPKPWI 47

## RESULT 5

US-10-051-852-11

; Sequence 11, Application US/10051852

; Patent No. US20020146411A1

; GENERAL INFORMATION:

; APPLICANT: Blackburn, Michael

; Church, William

; Gross, Mitchell

; Feuerstein, Giora

; Nichols, Andrew

; Padlan, Eduardo

; Patel, Arunbhai

; Sylvester, Daniel

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/051,852

FILING DATE: 17-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/344,050

FILING DATE: 25-JUN-1999

APPLICATION NUMBER: 08/783,853

FILING DATE: 16-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: P50438

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096

TELEFAX: <Unknown>

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-10-051-852-11  
Query Match 44.3%; Score 47; DB 12; Length 107;  
Best Local Similarity 100.0%; Pred. No. le-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QIVLSQSPAILSPGKVTMTCRASSVNYMHVYQKPGSSPKPWI 47

Db 1 QIVLSQSPAILSPGKVTMTCRASSVNYMHVYQKPGSSPKPWI 47

## RESULT 6

US-09-099-103

; Sequence 103, Application US/09965099

; Patent No. US20020136725A1

; GENERAL INFORMATION:

; APPLICANT: Blackburn, Michael

; Feuerstein, Giora

; Patel, Arunbhai

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN

TREATMENT OF THROMBOSIS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/965,099  
FILING DATE: 26-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/346,487  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
US-09-965-099-103

Query Match 44.3%; Score 47; DB 10; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.le-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47  
DB 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47

RESULT 7  
US-10-051-852-103  
Sequence 103, Application US/10051852  
Patent No. US20020146411A1  
GENERAL INFORMATION:  
APPLICANT: Blackburn, Michael  
Church, William  
Gross, Mitchell  
Feuerstein, Gloria  
Nichols, Andrew  
Padlan, Eduardo  
Patel, Arunbhai  
Sylvester, Daniel  
TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT  
OF THROMBOSIS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/051,852

FILING DATE: 17-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/344,050  
FILING DATE: 25-JUN-1999  
APPLICATION NUMBER: 08/783,853  
FILING DATE: 16-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: <Unknown>  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
US-10-051-852-103

Query Match 44.3%; Score 47; DB 12; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.le-34;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47  
DB 1 QIVLSQSPAILSPGKVTMTCRASSVNMHWYQKPGSSPKPW 47

RESULT 8  
US-09-982-107-12  
Sequence 12, Application US/09982107  
Patent No. US2002015958A1  
GENERAL INFORMATION:  
APPLICANT: HIATT, ANDREW C.  
APPLICANT: HEIN, MICHAEL  
TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOGLOBULINS CONTAINING  
PROTECTION PROTEINS IN PLANTS AND THEIR USE  
FILE REFERENCE: EPI3002E  
CURRENT APPLICATION NUMBER: US/09/982,107  
CURRENT FILING DATE: 2001-10-16  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 12  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Guy's 13 Kappa  
US-09-982-107-12

Query Match 25.5%; Score 27; DB 9; Length 105;  
Best Local Similarity 100.0%; Pred. No. 3.6e-17;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASGVPARFSGSGTSLTISR 76  
DB 50 TSNLASGVPARFSGSGTSLTISR 76

RESULT 9  
US-09-976-787-24  
Sequence 24, Application US/09976787  
Patent No. US20020064528A1

```

RESULT 11
US-09-144-886-88
; Sequence 88, Application US/09144886
; Patent No. US20020155114A1
; GENERAL INFORMATION:
; APPLICANT: Marks, James D
; APPLICANT: Amersdorfer, Peter
; TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize
; TITLE OF INVENTION: Botulinum Neurotoxins
; FILE REFERENCE: 2500.117USO
; CURRENT APPLICATION NUMBER: US/09/144.886
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88

```

Query Match 25.5%; Score 27; DB 10; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.7e-17;  
Matches 27; Conservative 0; Mismatches 0; Indels

Db 55 TSNLASGVPARFSGSGGTYSYLTISR 81  
Search completed: November 27, 2002, 07:32:53  
Job time : 9.82239 secs

QY 50 TSNLASGVPARFSGSGGTYSYLTISR 76  
|||||  
Db 50 TSNLASGVPARFSGSGGTYSYLTISR 76

## RESULT 14

US-09-910-059-9  
; Sequence 9, Application US/09910059  
; Patent No. US20020142359A1  
; GENERAL INFORMATION:  
; APPLICANT: Copley, Clive G  
; APPLICANT: Edge, Michael Derek  
; APPLICANT: Emery, Stephen Charles  
; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said Antibody,  
; TITLE OF INVENTION: Their Therapeutic use in an Adept System  
; FILE REFERENCE: 1991-209  
; CURRENT APPLICATION NUMBER: US/09/910,059  
; CURRENT FILING DATE: 2001-07-23  
; PRIOR APPLICATION NUMBER: US 09/171,945  
; PRIOR FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: PCT/GB97/01165  
; PRIOR FILING DATE: 1997-04-29  
; PRIOR APPLICATION NUMBER: GB 9703103.3  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: GB9609405.7  
; PRIOR FILING DATE: 1996-05-04  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-910-059-9

Query Match 25.5%; Score 27; DB 10; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.7e-17;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASGVPARFSGSGGTYSYLTISR 76  
|||||  
Db 50 TSNLASGVPARFSGSGGTYSYLTISR 76

## RESULT 15

US-09-144-886-89  
; Sequence 89, Application US/09144886  
; Patent No. US20020155114A1  
; GENERAL INFORMATION:  
; APPLICANT: Marks, James D  
; APPLICANT: Amersdorfer, Peter  
; TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize  
; TITLE OF INVENTION: Botulinum Neurotoxins  
; FILE REFERENCE: 2500.117USO  
; CURRENT APPLICATION NUMBER: US/09/144,886  
; CURRENT FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 98  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 89  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: BONT/A clone  
; OTHER INFORMATION: 3C3 region VL epitope 2  
US-09-144-886-89

Query Match 25.5%; Score 27; DB 9; Length 112;  
Best Local Similarity 100.0%; Pred. No. 3.8e-17;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TSNLASGVPARFSGSGGTYSYLTISR 76  
|||||

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:38:32 ; Search time 8.28185 Seconds  
(without alignments)  
241.342 Million cell updates/sec

Title: US-09-893-615-1

Perfect score: 91

Sequence: 1 WRMYFSRHAHLRSP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	15	AAW94701	Lipoteichoic acid
2	91	100.0	19	AAW94733	Sequence 15mer 2nd
3	91	100.0	19	AAW94714	Sequence 15mer2-12
4	91	100.0	19	AAW94715	Sequence 15mer2-13
5	91	100.0	19	AAW94717	Sequence 15mer2-10
6	91	100.0	19	AAW94718	Sequence 15mer2-16
7	91	100.0	19	AAW94719	Sequence 15mer2-17
8	66	72.5	19	AAW94716	Sequence 15mer2-14
9	47	51.6	757	20 AAY34472	Porphorymonas ging
10	47	51.6	763	20 AAY34348	Porphorymonas ging

11	45	49.5	117	22	AAO10712	Human polypeptide
12	45	49.5	684	22	ABG62691	Drosophila melanog
13	45	49.5	684	23	AAU74625	Oestrogen-regulate
14	43	47.3	82	22	AAU52691	Propionibacterium
15	43	47.3	88	22	AAU41763	Propionibacterium
16	43	47.3	95	19	AAU21419	Human high mobil
17	43	47.3	860	22	AAU88351	Human membrane or
18	43	47.3	1358	19	AAW57837	Aldehyde oxidase p
19	43	47.3	1358	19	AAW57838	Aldehyde oxidase p
20	42	46.2	147	23	AAU48365	Insulin-like growt
21	42	46.2	252	23	ABU49295	Listeria monocytog
22	42	46.2	486	22	ABU59435	Drosophila melanog
23	42	46.2	2785	21	AAV57148	Human down-regulat
24	41.5	45.6	248	22	ABU53062	Escherichia coli p
25	41	45.1	127	22	AAU60591	Human brain expres
26	41	45.1	127	22	AAU73259	Human bone marrow
27	41	45.1	127	22	AAU33460	Peptide #7497 enco
28	41	45.1	127	23	ABG43106	Human peptide enco
29	41	45.1	191	23	AAU69540	Human G protein-co
30	41	45.1	196	22	ABG26403	Novel human diago
31	41	45.1	221	22	ABG05725	Novel human diago
32	41	45.1	230	22	AAU93340	Human polypeptide,
33	41	45.1	344	22	AAU65923	Propionibacterium
34	41	45.1	428	22	AAU64689	Propionibacterium
35	41	45.1	434	22	AAU96758	Putative P. abyssi
36	41	45.1	452	22	AAU94445	Human protein sequ
37	41	45.1	519	22	ABG09372	Novel human diago
38	41	45.1	651	21	AAU29652	Human membrane-ass
39	41	45.1	680	23	AAE19421	Human furin mutant
40	41	45.1	809	22	ABG09902	Novel human diago
41	41	45.1	1014	22	AAU40124	Human polypeptide
42	41	45.1	1063	23	ABU90934	Herbicidally activ
43	41	45.1	1066	22	AAU41910	Human polypeptide
44	41	45.1	1069	21	AAU42717	Human ORFX ORF2481
45	41	45.1	1069	22	ABG09904	Novel human diago

ALIGNMENTS

RESULT 1

AAW94701

ID AAW94701 standard; peptide; 15 AA.

XX AAW94701;

XX 22-APR-1999 (first entry)

XX Lipoteichoic acid epitope peptide mimic for Mab 96-110.

XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;

XX immunoglobulin; phagocytosis; infection; epitope; peptide mimic;

XX Mab 96-110.

XX Staphylococcus sp.

XX WO9857994-A2.

XX 23-DEC-1998.

XX 16-JUN-1998; 98WO-US12402.

XX 16-JUN-1997; 97US-0049871.

XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

XX Fischer GW, Schuman RF, Stinson JL, Wong H;

XX WPI; 1999-095329/08.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
PT to develop products for the diagnosis, prevention and treatment of  
PT infections caused by gram positive bacteria

XX Claim 16; Page 120; 150pp; English.  
 PS The invention relates to a monoclonal antibody (Mab) to lipoteichoic acid  
 CC of gram positive bacteria, where the Mab is a chimeric immunoglobulin  
 CC comprising at least part of a human immunoglobulin constant region and  
 CC at least part of a non-human immunoglobulin variable region having  
 CC specificity to lipoteichoic acid of gram positive bacteria. The  
 CC antibodies bind to whole bacteria and enhance phagocytosis and killing of  
 CC the bacteria and enhance protection from lethal infection. The antibodies  
 CC or peptides (encoded by a DNA of the variable region of anti-lipoteichoic  
 CC acid antibody or characterised by amino acids corresponding to one or  
 CC more of the complementarity determining regions (CDRs) of the variable  
 CC region of the antibody) can be used for treating or preventing infections  
 CC caused by gram positive bacteria. They can also be used for the diagnosis  
 CC of gram positive bacterial infections. The present sequence represents a  
 CC specifically claimed lipoteichoic acid epitope peptide mimic that can be  
 CC bound by the antibody of the invention (Mab 96-110).  
 XX  
 SQ Sequence 15 AA;

Query Match 100.0%; Score 91; DB 20; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
 ID AAW94733  
 XX  
 AC AAW94733;  
 DT 22-APR-1999 (first entry)  
 XX

DE Sequence 15mer 2nd.12 resulting from library panning experiments.  
 XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW Mab 96-110; panning.  
 XX Staphylococcus sp.  
 OS WO9857994-A2.  
 PN 23-DEC-1998.  
 PD 16-JUN-1998; 98WO-US12402.  
 PF 16-JUN-1997; 97US-0049871.  
 PR (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
 XX Fischer GW, Schuman RF, Stinson JL, Wong H;  
 PI WPI: 1999-095329/08.  
 DR N-PSDB; AAX05359.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 XX Example 5; Fig 8; 150pp; English.  
 PS The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and

CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94726-34 represent common peptide sequences resulting from  
 CC all library panning experiments. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
 XX

SQ Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
 ID AAW94714  
 XX  
 AC AAW94714;  
 DT 22-APR-1999 (first entry)  
 XX

DE Sequence 15mer2-12/0 resulting from 15mer library panning experiment.  
 XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW Mab 96-110; panning.  
 XX Staphylococcus sp.  
 OS WO9857994-A2.  
 PN 23-DEC-1998.  
 PD 16-JUN-1998; 98WO-US12402.  
 PF 16-JUN-1997; 97US-0049871.  
 PR (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
 XX Fischer GW, Schuman RF, Stinson JL, Wong H;  
 PI WPI: 1999-095329/08.  
 DR N-PSDB; AAX05540.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 XX Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also



CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
 XX

SO Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 8.4e-09; Mismatches 0; Gaps 0;  
 Matches 15; Conservative 0; Indels 0;

Oy 1 WRMYFSHRHAHLRSP 15

Db 3 WRMYFSHRHAHLRSP 17

RESULT 4

AAW94715  
 ID AAW94715 standard; Protein: 19 AA.

AC AAW94715;

DT 22-APR-1999 (first entry)

XX Sequence 15mer2-13/0 resulting from 15mer library panning experiment.

XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW Mab 96-110; panning.

XX Staphylococcus sp.  
 OS WO9857994-A2.

XX 23-DEC-1998.

XX 16-JUN-1998; 98WO-US12402.

XX 16-JUN-1997; 97US-0049871.

XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

XX Fischer GW, Schuman RF, Stinson JL, Wong H;

XX WPI; 1999-095329/08.

XX N-PSDB; AAX05541.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria

PS Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 8.4e-09; Mismatches 0; Gaps 0;  
 Matches 15; Conservative 0; Indels 0;

Oy 1 WRMYFSHRHAHLRSP 15

Db 3 WRMYFSHRHAHLRSP 17

RESULT 5

AAW94717  
 ID AAW94717 standard; Protein: 19 AA.

AC AAW94717;

XX 22-APR-1999 (first entry)

XX Sequence 15mer2-10/0 resulting from 15mer library panning experiment.

XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW Mab 96-110; panning.

XX Staphylococcus sp.  
 OS WO9857994-A2.

XX 23-DEC-1998.

XX 16-JUN-1998; 98WO-US12402.

XX 16-JUN-1997; 97US-0049871.

XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

XX Fischer GW, Schuman RF, Stinson JL, Wong H;

XX WPI; 1999-095329/08.

XX N-PSDB; AAX05543.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria

PS Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 8.4e-09; Mismatches 0; Gaps 0;  
 Matches 15; Conservative 0; Indels 0;

QY 1 WRMYFSHRHAHLRSP 15  
 ID |||||  
 DB 3 WRMYFSHRHAHLRSP 17

## RESULT 6

AAW94718  
 ID AAW94718 standard; Protein; 19 AA.

XX AC AAW94718;

XX DT 22-APR-1999 (first entry)

XX Sequence 15mer2-16/0 resulting from 15mer library panning experiment.

XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW Mab 96-110; panning.

XX OS Staphylococcus sp.

XX PN WO9857994-A2.

XX PD 23-DEC-1998.

XX PF 16-JUN-1998; 98WO-US12402.

XX PR 16-JUN-1997; 97US-0049871.

XX PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

XX PI Fischer GW, Schuman RF, Stinson JL, Wong H;

XX DR WPI: 1999-095329/08.

XX DR N-PSDB; AAX05544.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria

XX Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.

CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 8.4e-09;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15

ID |||||

DB 3 WRMYFSHRHAHLRSP 17

## RESULT 7

AAW94719  
 ID AAW94719 standard; Protein; 19 AA.

XX AC AAW94719;

XX DT 22-APR-1999 (first entry)

XX Sequence 15mer2-17/0 resulting from 15mer library panning experiment.

XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW Mab 96-110; panning.

XX OS Staphylococcus sp.

XX PN WO9857994-A2.

XX PD 23-DEC-1998.

XX PF 16-JUN-1998; 98WO-US12402.

XX PR 16-JUN-1997; 97US-0049871.

XX PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

XX PI Fischer GW, Schuman RF, Stinson JL, Wong H;

XX DR WPI: 1999-095329/08.

XX DR N-PSDB; AAX05545.

XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria

XX Example 5; Fig 6A-B; 150pp; English.

XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.

CC Sequences AAW94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.

XX Sequence 19 AA;

Query Match 100.0%; Score 91; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 8.4e-09;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15

ID |||||

DB 3 WRMYFSHRHAHLRSP 17

## RESULT 8

AAW94716  
 ID AAW94716 standard; Protein; 19 AA.

XX AC AAW94716;

XX

DT 22-APR-1999 (first entry)  
 DE Sequence 15mer2-14/0 resulting from 15mer library panning experiment.  
 XX Monoclonal antibody; Mab; lipoteichoic acid; gram positive; bacteria;  
 KW immunoglobulin; phagocytosis; infection; epitope; peptide mimic;  
 KW Mab 96-110; panning.  
 XX Staphylococcus sp.  
 OS  
 XX WO9857994-A2.  
 PN 23-DEC-1998.  
 XX 16-JUN-1998; 98WO-US12402.  
 XX 16-JUN-1997; 97US-0049871.  
 PR (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
 XX Fischer GW, Schuman RF, Stinson JL, Wong H;  
 PI WPI; 1999-095329/08.  
 DR N-PSDB; AAX05542.  
 XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 PT to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 XX Example 5; Fig 6A-B; 150pp; English.  
 PS The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 XX acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAM94705-22 represent sequences resulting from a 15mer library  
 CC panning second experiment. Three series of panning experiments were  
 CC conducted to identify peptide sequences to which antibody of the  
 CC invention (Mab 96-110) bound strongly. The translated sequences provide  
 CC lipoteichoic acid epitope peptide mimics to which Mab96-110 bound.  
 XX Sequence 19 AA;  
 SQ  
 Query Match 72.5%; Score 66; DB 20; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.00015;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Oy 1 WRMYFSHRHAHLRSP 15  
 II III: IIII II  
 Db 3 WRKYFSYHHAHLCSF 17  
 RESULT 9  
 AAY34472  
 ID AAY34472 standard; Protein; 757 AA.  
 XX  
 AC AAY34472;  
 XX 25-AUG-1999 (first entry)  
 DT Porphorymonas gingivalis protein PG13.  
 DE  
 XX Porphorymonas gingivalis; PG; periodontal disease; gingivitis;  
 KW vaccine; antigenic.

XX Porphorymonas gingivalis.  
 OS  
 XX WO9929870-A1.  
 PN 17-JUN-1999.  
 XX 10-DEC-1998; 98WO-AU01023.  
 XX 04-AUG-1998; 98AU-0005028.  
 PR 10-DEC-1997; 97AU-0000839.  
 PR 31-DEC-1997; 97AU-0001182.  
 PR 30-JAN-1998; 98AU-0001546.  
 PR 10-MAR-1998; 98AU-0002264.  
 PR 09-APR-1998; 98AU-0002911.  
 PR 23-APR-1998; 98AU-0003128.  
 PR 05-MAY-1998; 98AU-0003338.  
 PR 22-MAY-1998; 98AU-0003654.  
 PR 29-JUL-1998; 98AU-0004917.  
 XX (CSLC-) CSL LTD.  
 PA Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;  
 PI Ross BC, Rothel LJ, Webb EA;  
 XX WPI; 1999-385613/32.  
 DR N-PSDB; AAX91690.  
 XX Antigenic Porphorymonas gingivalis peptides for preventing  
 PT gingivitis  
 PT Claim 1; Page 450-451; 588pp; English.  
 PS AX91536 to AX91801 encode two hundred and sixty six antigenic  
 CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to  
 CC AX91583. AX91802 to AX91989 represent PCR primers used in the  
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial  
 CC activity with a vaccine mechanism of action. The PG polypeptides can be  
 CC used as vaccines especially against Porphorymonas gingivalis. Probes can  
 CC be used to detect Porphorymonas gingivalis in standard hybridisation  
 CC assays. Porphorymonas gingivalis is involved in periodontal disease  
 CC especially gingivitis.  
 XX Sequence 757 AA;  
 SQ  
 Query Match 51.6%; Score 47; DB 20; Length 757;  
 Best Local Similarity 61.5%; Pred. No. 17;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Oy 1 WRMYFSHRHAHLR 13  
 II: IIII II  
 Db 700 WRLATSHRFAHER 712  
 RESULT 10  
 AAY34348  
 ID AAY34348 standard; Protein; 763 AA.  
 XX  
 AC AAY34348;  
 XX 25-AUG-1999 (first entry)  
 DT Porphorymonas gingivalis protein PG13.  
 DE  
 XX Porphorymonas gingivalis; PG; periodontal disease; gingivitis;  
 KW vaccine; antigenic.  
 XX Porphorymonas gingivalis.  
 OS  
 XX WO9929870-A1.  
 PN 17-JUN-1999.  
 XX

PF 10-DEC-1998; 98WO-AU01023.  
 XX  
 PR 04-AUG-1998; 98AU-0005028.  
 PR 10-DEC-1997; 97AU-0000839.  
 PR 31-DEC-1997; 97AU-0001182.  
 PR 30-JAN-1998; 98AU-0001546.  
 PR 10-MAR-1998; 98AU-0002264.  
 PR 09-APR-1998; 98AU-0002911.  
 PR 23-APR-1998; 98AU-0003128.  
 PR 05-MAY-1998; 98AU-0003338.  
 PR 22-MAY-1998; 98AU-0003654.  
 PR 29-JUL-1998; 98AU-0004917.  
 XX

(CSLC-) CSL LTD.

PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;  
 PI Ross BC, Rothel LJ, Webb EA;

DR WPI; 1999-385613/32.  
 DR N-PSDB; AAX91566.

XX Antigenic Porphyromonas gingivalis peptides for preventing  
 PT gingivitis

PS Claim 1; Page 308-309; 588pp; English.

XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic  
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAX34318 to  
 CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the  
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial  
 CC activity with a vaccine mechanism of action. The PG polypeptides can be  
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can  
 CC be used to detect Porphyromonas gingivalis in standard hybridisation  
 CC assays. Porphyromonas gingivalis is involved in periodontal disease  
 CC especially gingivitis.

XX Sequence 763 AA;

Query Match 51.6%; Score 47; DB 20; Length 763;  
 Best Local Similarity 61.5%; Pred. No. 17;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLR 13  
 II: III III I  
 DB 706 WRLATSHRFAHR 718

RESULT 11  
 AAO10712  
 ID AAO10712 standard; Protein; 117 AA.

XX AAO10712;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 24604.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

XX 07-SEP-2001.

PD 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR N-PSDB; AAI90843.

XX Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders

XX Claim 20; SEQ ID NO 24604; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 117 AA;

Query Match 49.5%; Score 45; DB 22; Length 117;

Best Local Similarity 66.7%; Pred. No. 4.5;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 YFSHRHAHLRSP 15

II: III III II

DB 71 YRSHTHAHTSP 82

RESULT 12

ABB62691  
 ID ABB62691 standard; Protein; 684 AA.

XX ABB62691;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 14865.

KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL06794.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions

XX Disclosure; SEQ ID NO 14865; 21pp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent.  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABBS7737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 684 AA;  
 SQ

Query Match 49.5%; Score 45; DB 22; Length 684;  
 Best Local Similarity 70.0%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 SHRAHLRSP 15  
 DB 505 SHRHGHVSP 514  
 |||||: ||  
 |||||: ||

RESULT 13  
 AAU74625  
 ID AAU74625 standard; Protein; 684 AA.  
 XX  
 AC AAU74625;  
 XX  
 XX 09-APR-2002 (first entry)  
 DT  
 XX Oestrogen-regulated LIV-1 family protein AAF50401\_Dm.  
 DE  
 XX LIV-1; oestrogen; cytostatic; neuroprotective; zinc homeostasis;  
 KW gene therapy; apoptosis modulator; cancer; neurodegenerative disorder;  
 KW apoptotic disorder; zinc-homeostasis related disorder.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200196372-A2.  
 XX  
 PD 20-DEC-2001.  
 XX  
 XX 13-JUN-2001; 2001WO-GB02597.  
 PF  
 XX 13-JUN-2000; 2000GB-0014411.  
 PR  
 PR 14-JUN-2000; 2000GB-0014493.  
 PR  
 XX 05-JUL-2000; 2000US-216349P.  
 XX  
 PA (UYCA-) UNIV COLLEGE CARDIFF.  
 XX  
 XX Taylor KM, Morgan HE, Nicholson RI;  
 PI  
 XX WPI; 2002-106465/14.  
 DR  
 XX Use of a polypeptide comprising one or more consensus regions of  
 PT proteins of LIV-1 family for treating disorders of zinc homeostasis  
 PT such as breast cancer, neurodegenerative disorders, and for modifying  
 PT apoptosis -  
 XX  
 XX Claim 1; Figure 1; 67pp; English.  
 PS  
 XX The invention describes the a polypeptide comprising one or more  
 CC consensus regions of proteins of LIV-1 family or its functional  
 CC homologue. The polypeptide is useful in the preparation of a medicament  
 CC for the treating a disease e.g. those involving disorders of zinc  
 CC homeostasis, in gene therapy and for modifying apoptosis in vitro or in  
 CC vivo on contact with cells. Diseases involving defects in zinc  
 CC homeostasis include cancer, neurodegenerative disorders and apoptotic  
 CC disorders. Recombinant proteins of the LIV-1 family (an

CC oestrogen-regulated gene) are useful for diagnosing a zinc  
 CC homeostasis-related condition in a subject which involves contacting a  
 CC sample from the subject with the recombinant protein and measuring the  
 CC binding of antibody to the sample. The antibody is also useful for  
 CC treating a zinc homeostasis-related condition. This sequence is a  
 CC member of the LIV-1 family (a gene regulated by oestrogen levels) and  
 CC is useful for creating recombinant proteins for diagnosing  
 CC zinc-homeostasis related conditions, described in the method of the  
 CC invention.  
 XX  
 SQ Sequence 684 AA;  
 Query Match 49.5%; Score 45; DB 23; Length 684;  
 Best Local Similarity 70.0%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 SHRAHLRSP 15  
 DB 505 SHRHGHVSP 514  
 |||||: ||  
 |||||: ||

RESULT 14  
 AAU52691  
 ID AAU52691 standard; Protein; 82 AA.  
 XX  
 AC AAU52691;  
 XX  
 XX 27-FEB-2002 (first entry)  
 DT  
 XX  
 DE Proplonibacterium acnes immunogenic protein #13587.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Proplonibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 XX 20-APR-2001; 2001WO-US12865.  
 PF  
 XX 21-APR-2000; 2000US-199047P.  
 PR  
 PR 02-JUN-2000; 2000US-208841P.  
 PR  
 XX 07-JUL-2000; 2000US-216747P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 DR WPI; 2001-616774/71.  
 DR N-PSDB; AAS59556.  
 XX  
 XX Proplonibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 XX  
 PS Example 1; SEQ ID No 13886; 1069pp; English.  
 XX  
 XX Sequences AAU39105-AAU68017 represent Proplonibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 82 AA;

Query Match 47.3%; Score 43; DB 22; Length 82;

Best Local Similarity 70.0%; Pred. No. 6.6;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WRMYFSRHA 10

Db 35 WRWRWSRHA 44

RESULT 15

AAU41763

ID AAU41763 standard; Protein; 88 AA.

XX

AC AAU41763;

XX

DT 13-FEB-2002 (first entry)

XX

DE Propionibacterium acnes immunogenic protein #2659.

XX

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.

XX

OS Propionibacterium acnes.

XX

FN WO200181581-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12865.

XX

PR 21-APR-2000; 2000US-199047P.

PR

PR 02-JUN-2000; 2000US-208841P.

XX

PR 07-JUL-2000; 2000US-216747P.

XX

PA (CORI-) CORIXA CORP.

XX

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhátia A;

PI

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX

WPI: 2001-616774/71.

DR

DR N-PSDB; AAS59515.

XX

PT Propionibacterium acnes polypeptides and nucleic acids useful for

XX

PT vaccinating against and diagnosing infections, especially useful for

XX

PT treating acne vulgaris .

XX

PS Example 1; SEQ ID No 2958; 1069pp; English.

XX

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 88 AA;

Query Match 47.3%; Score 43; DB 22; Length 88;

Best Local Similarity 72.7%; Pred. No. 7.1;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 YFSHRHAHLRS 14

Db 7 YVHRRHARLRS 17

Search completed: November 27, 2002, 07:17:47

Job time : 19.2819 secs

GenCore version 5.1.3  
Copyright (C) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:15:05 ; Search time 4.11197 Seconds  
(without alignments)  
350.688 Million cell updates/sec

Title: US-09-893-615-1

Perfect score: 91

Sequence: 1 WRMYFSHRHAHLRSP 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

PIR-73:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Ident				
1	45	49.5	215	2	AT0355	probable membrane	
2	45	49.5	2	2	T48357	hypothetical prote	
3	44	48.4	307	2	F69898	hypothetical prote	
4	44	48.4	703	2	T04191	hypothetical prote	
5	44	48.4	743	2	T02147	hypothetical prote	
6	43	47.3	424	2	JC5891	omega 6 desaturase	
7	43	47.3	1358	2	T01698	aldehyde oxidase (	
8	42	46.2	134	2	C69849	hypothetical prote	
9	42	46.2	252	2	AC1181	hypothetical prote	
10	42	46.2	252	2	AD1538	hypothetical prote	
11	42	46.2	929	2	T29162	hypothetical prote	
12	41.5	45.6	248	2	JQ1340	trax protein - Esc	
13	41	45.1	208	2	AE2378	hypothetical prote	
14	41	45.1	241	2	G69899	transcription regu	
15	41	45.1	257	1	B64042	formate dehydrogen	
16	41	45.1	286	2	E83048	hypothetical prote	
17	41	45.1	346	2	S27482	hypothetical prote	
18	41	45.1	370	2	F87456	transcription regu	
19	41	45.1	434	2	D75212	asparagine-tRNA li	
20	41	45.1	470	2	A22155	hypothetical prote	
21	41	45.1	822	2	T01095	hypothetical prote	
22	41	45.1	838	2	S61250	DNA helicase/prima	
23	41	45.1	1063	2	T00624	endo-1,4-beta-xyla	
24	40	44.0	167	2	B82636	hypothetical prote	
25	40	44.0	321	2	T48469	hypothetical prote	
26	40	44.0	349	2	T03500	hypothetical prote	
27	40	44.0	351	1	A48763	transcription fact	
28	40	44.0	431	2	A64658	LPS biosynthesis p	
29	40	44.0	433	2	D75632	probable hemolysin	

30	40	44.0	555	2	AE2147	cytochrome c oxida
31	40	44.0	710	2	T20635	hypothetical prote
32	40	44.0	838	2	A48410	ring-infected eryt
33	40	44.0	899	2	H87513	hypothetical prote
34	40	44.0	1349	2	T01699	aldehyde oxidase (
35	40	44.0	1495	2	S60255	transcription co-r
36	39.5	43.4	1167	2	T42704	hypothetical prote
37	39	42.9	95	2	S43068	hypothetical prote
38	39	42.9	296	2	A40996	phenylalanine 4-mo
39	39	42.9	303	2	E95374	probable ABC trans
40	39	42.9	323	2	E95877	probable ABC trans
41	39	42.9	349	2	S43770	phosphatidylcholin
42	39	42.9	351	2	S11519	phosphatidylcholin
43	39	42.9	508	2	E83788	galactose-1-phosph
44	39	42.9	548	2	T40508	probable asparagin
45	39	42.9	562	2	E70399	hydrogenase regula

#### ALIGNMENTS

##### RESULT 1

AI0355

probable membrane protein YP02924 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C:Accession: AI0355

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, G

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G

ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AI0355

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC92172.1; PID:g15980886; GSPDB:GN00175

C:Genetics:

A:Gene: YP02924

Query Match 49.5%; Score 45; DB 2; Length 215;

Best Local Similarity 61.5%; Pred. No. 4.7;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WRMYFSHRHAHLR 13

Db 72 WAMTGRREAHLR 84

##### RESULT 2

T48357

hypothetical protein F12B4.90 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T48357

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke,

submitted to the Protein Sequence Database, March 2000

A:Reference number: 224492

A:Accession: T48357

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-684 <BEV>

A:Cross-references: EMBL:AL162751

A:Experimental source: cultivar Columbia; BAC clone F12E4

C:Genetics:

A:Map position: 5

A:Note: F12B4.90

Query Match 49.5%; Score 45; DB 2; Length 684;

Best Local Similarity 70.0%; Pred. No. 17;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 YFSHRHAHLR 13  
 :|||: |||  
 Db 412 HFSKHHLR 421  
 :|||: |||  
 RESULT 3  
 hypothetical protein yobF - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C:Accession: F69898  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
 C:Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
 A:Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
 lech, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 keuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tojnoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: F69898  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-307 <KUN>  
 A:Cross-references: GB:Z99114; GB:AL009126; NID:G2634230; PIDN:CAB13781.1; PID:G2634282  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yobF  
 C:Superfamily: Bacillus subtilis hypothetical protein yobF  
 Query Match 48.4%; Score 44; DB 2; Length 307;  
 Best Local Similarity 58.3%; Pred. No. 10;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 WRMYFSHRHAHL 12  
 :|||: |||  
 Db 247 YAMYKDRSHL 258  
 :|||: |||  
 RESULT 4  
 hypothetical protein T4F9.20 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999  
 C:Accession: T04191  
 R:Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.  
 submitted to the Protein Sequence Database, March 1999  
 A:Reference number: Z15260  
 A:Accession: T04191  
 A:Molecule type: DNA  
 A:Residues: 1-703 <BEV>  
 A:Cross-references: EMBL:AL049523  
 A:Experimental source: cultivar Columbia; BAC clone T4F9  
 C:Genetics:  
 A:Map position: 4  
 A:Note: T4F9.20  
 Query Match 48.4%; Score 44; DB 2; Length 703;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 YFSHRHAHLR 13  
 :|||: |||  
 Db 412 HFSKHHLR 421  
 :|||: |||

RESULT 5  
 T02147  
 hypothetical protein F8K4.22 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 22-Oct-1999  
 C:Accession: T02147  
 R:Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Kwan, A.; Yu, G.; Oji, O.; Liu, S.;  
 rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel  
 submitted to the EMBL Data Library, August 1998  
 A:Description: Arabidopsis thaliana chromosome 1 BAC F8K4 sequence.  
 A:Reference number: Z14574  
 A:Accession: T02147  
 A:Status: translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-743 <VYS>  
 A:Cross-references: EMBL:AC004392; NID:G3282170; PID:G3367535; GSPDB:CN00059; ATSP:F8  
 C:Genetics:  
 A:Gene: ATSP:F8K4.22  
 A:Map position: 1  
 Query Match 48.4%; Score 44; DB 2; Length 743;  
 Best Local Similarity 70.0%; Pred. No. 27;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 YFSHRHAHLR 13  
 :|||: |||  
 Db 418 HFSKHHLR 427  
 :|||: |||  
 RESULT 6  
 JCS891  
 omega 6 desaturase (EC 1.14.99.-) precursor, chloroplast - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 18-Mar-1998 #sequence\_revision 18-Mar-1998 #text\_change 20-Jun-2000  
 C:Accession: JCS891  
 R:Sato, N.; Fujiwara, S.; Kawaguchi, A.; Tsuzuki, M.  
 J. Biochem. 122, 1224-1232, 1997  
 A:Title: Cloning of a gene for chloroplast omega 6 desaturase of a green alga, Chlamy  
 A:Reference number: JCS891; MUID:98158334; PMID:9498569  
 A:Accession: JCS891  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-424 <SAT>  
 A:Cross-references: DDBJ:AB007640; NID:G2696716; PIDN:BAA23881.1; PID:G2696717  
 C:Comment: This enzyme catalyzes the desaturation of monoenoic to dienoic acids.  
 C:Genetics:  
 A:Gene: des6  
 C:Superfamily: omega-3 fatty acid desaturase  
 C:Keywords: chloroplast; oxidoreductase  
 F:1-40/Domain: transit peptide (chloroplast) #status predicted <TPS>  
 Query Match 47.3%; Score 43; DB 2; Length 424;  
 Best Local Similarity 54.5%; Pred. No. 21;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 WRMYFSHRHAHL 11  
 :|||: |||  
 Db 178 WRKHNHHHAHL 188  
 :|||: |||  
 RESULT 7  
 T01698  
 aldehyde oxidase (EC 1.2.3.1) - maize  
 C:Species: Zea mays (maize)  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 28-Jul-2000  
 C:Accession: T01698  
 R:Sekimoto, H.; Seo, M.; Dohmae, N.; Takio, K.; Kamiya, Y.; Koshiba, T.  
 J. Biol. Chem. 272, 15280-15285, 1997  
 A:Title: Cloning and molecular characterization of plant aldehyde oxidase.  
 A:Reference number: Z14401; MUID:97326103; PMID:9182554  
 A:Accession: T01698  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA



A:Residues: 1-1358 <SEK>  
A:Cross-references: EMBL:D88451; NID:g2589161; PIDN:BAA23226.1; PID:g2589162  
C:Genetics:  
A:Gene: AO-1  
C:Function:  
A:Description: catalyzes oxidation of aldehydes into respective acids  
C:Superfamily: xanthine dehydrogenase; ferredoxin (2Fe-2S) homology  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase  
F:33-81/Domain: ferredoxin (2Fe-2S) homology <PDX>  
F:50.55.58.80/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 47.3%; Score 43; DB 2; Length 1358;  
Best Local Similarity 58.7%; Pred. No. 75;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 MYFSHRHAHLRSP 14  
: : : : :  
Db 619 IYSTHPAHVRSP 630

## RESULT 8

C69849

hypothetical protein yjdh - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000

C:Accession: C69849

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choc, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Laubert, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PID:9384377

A:Accession: C69849

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-134 <KUN>

A:Cross-references: GB:299110; GB:AL009126; NID:g2633472; PIDN:CAB13062.1; PID:g2633559

A:Experimental source: strain 168

C:Genetics:

A:Gene: yjdh

C:Superfamily: Bacillus subtilis hypothetical protein yjdh

Query Match 46.2%; Score 42; DB 2; Length 134;  
Best Local Similarity 66.7%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 YFSHRHAHLRSP 15  
: : : : :  
Db 21 YFSHTNPQRSP 32

## RESULT 9

AC1181

hypothetical protein lmo0851 [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes

C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AC1181

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1181

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-252 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAC98929.1; PID:g16410254; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo0851

Query Match 46.2%; Score 42; DB 2; Length 252;  
Best Local Similarity 53.8%; Pred. No. 17;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MYFSHRHAHLRSP 15  
: : : : :  
Db 103 LYLAHTGRHLRSP 115

## RESULT 10

AD1538

hypothetical protein lin0844 [imported] - Listeria innocua (strain Clip11262)

C:Species: Listeria innocua

C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AD1538

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloek, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla, A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AD1538

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-252 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC96076.1; PID:g16413295; GSPDB:GN00178

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lin0844

Query Match 46.2%; Score 42; DB 2; Length 252;  
Best Local Similarity 53.8%; Pred. No. 17;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MYFSHRHAHLRSP 15  
: : : : :  
Db 103 LYLAHTGRHLRSP 115

## RESULT 11

T29162

hypothetical protein C01B7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T29162

R:Woessner, J.; Bradshaw, H.

submitted to the EMBL Data Library, March 1996

A:Description: The sequence of C. elegans cosmid C01B7.

A:Reference number: 220581

A:Accession: T29162

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-929 <WOB>

A:Cross-references: EMBL:U53147; PIDN:AAA96114.1; GSPDB:GN00023; CESP:C01B7.1

A:Experimental source: strain Bristol N2; clone C01B7

C:Genetics:

A:Gene: CESP:C01B7.1

A:Map position: 5

A:Introns: 288/1; 447/3; 478/3; 514/3; 541/1; 567/3; 883/2

Query Match 46.2%; Score 42; DB 2; Length 929;  
 Best Local Similarity 46.2%; Pred. No. 72;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 MYFSHRHAHLRSP 15  
 :: :||| | |  
 DB 453 LFVAHRHGRASP 465

## RESULT 12

JQ1340  
 trax protein - Escherichia coli plasmids  
 C:Species: Escherichia coli  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 08-Oct-1999  
 C:Accession: JQ1340; S10661  
 R:Cran, D.S.; Loh, S.M.; Cheah, K.C.; Skurray, R.A.  
 Gene 104, 85-90, 1991  
 A:Title: Sequence and conservation of genes at the distal end of the transfer region on  
 A:Reference number: JQ1338; MUID:92009201; PMID:1916281  
 A:Accession: JQ1340  
 A:Molecule type: DNA  
 A:Residues: 1-248 <CRA>  
 A:Cross-references: GB:M38048; NID:gl51835; PIDN:AAA98314.1; PID:gl51836  
 A:Experimental source: plasmid R6-5  
 R:Yoshioka, Y.; Fujita, Y.; Ohtsubo, E.  
 J. Mol. Biol. 214, 39-53, 1990  
 A:Title: Nucleotide sequence of the promoter-distal region of the tra operon of plasmid  
 A:Reference number: S10658; MUID:90317835; PMID:2164585  
 A:Accession: S10661  
 A:Molecule type: DNA  
 A:Residues: 1-248 <YOS>  
 A:Cross-references: EMBL:X55815; NID:g42620; PIDN:CAA39339.1; PID:g42625  
 A:Experimental source: plasmid R100  
 C:Genetics:  
 A:Gene: trax  
 A:Genome: plasmid  
 C:Keywords: transmembrane protein  
 F:39-55/Domain: transmembrane #status predicted <TM1>  
 F:180-196/Domain: transmembrane #status predicted <TM2>  
 F:200-216/Domain: transmembrane #status predicted <TM3>  
 F:232-248/Domain: transmembrane #status predicted <TM4>

Query Match 45.6%; Score 41.5; DB 2; Length 248;  
 Best Local Similarity 53.3%; Pred. No. 21;  
 Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 WRMYFSHRHAHLRSP 15  
 | : | ||| | |  
 DB 75 WGLNLS-RHAHRRP 88

## RESULT 13

AE2378  
 hypothetical protein all4581 [Imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AE2378  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AE2378  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-208 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BAB76280.1; PID:gi17133717; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all4581

Query Match 45.1%; Score 41; DB 2; Length 208;

Best Local Similarity 52.9%; Pred. No. 21;  
 Matches 9; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 1 WRMYFSH--RHAHLRSP 15  
 ||| | : | ||| | |  
 DB 81 WRMRSSNPIRRHAHRRP 97

## RESULT 14

G59899  
 transcription regulator AraC/XylS family homolog yobQ - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C:Accession: G69899  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
 C.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal  
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, S.;  
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Teynoni, A.; Tosato, V.; Uchiya  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtili  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: G69899  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-241 <KUN>  
 A:Cross-references: GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13797.1; PID:ell1853  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yobQ

Query Match 45.1%; Score 41; DB 2; Length 241;  
 Best Local Similarity 60.0%; Pred. No. 24;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RMVFSHRHAH 11  
 | : | ||| | |  
 DB 18 RLYHSHKHAY 27

## RESULT 15

B64042  
 formate dehydrogenase (EC 1.2.1.2) O gamma chain - Haemophilus influenzae (strain Rd  
 C:Species: Haemophilus influenzae  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: B64042  
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage  
 Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glöck, A.; Kelley, J.M.; Weidman  
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente  
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630; PMID:7542800  
 A:Accession: B64042  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-257 <TIGR>  
 A:Cross-references: GB:L42023; TIGR:HI0008  
 C:Superfamily: formate dehydrogenase gamma chain  
 C:Keywords: NAD; oxidoreductase

Query Match 45.1%; Score 41; DB 1; Length 257;  
 Best Local Similarity 85.7%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 WRMYFESH 7  
    |||  
Db 159 WRQYFESH 165

Search completed: November 27, 2002, 07:24:07  
Job time : 12.112 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:39:34 ; Search time 2.14286 Seconds  
(without alignments)  
290.334 Million cell updates/sec

Title: US-09-893-615-1

Perfect score: 91

Sequence: 1 WRMYFSRHAHLRSP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	45	49.5	724	1	CCT1_MOUSE		Q9qvv9 mus musculus
2	42	46.2	272	1	TNFS_CHICK		Q91868 gallus gall
3	41.5	45.6	248	1	TRX2_ECOLI		P22710 escherichia
4	41	45.1	238	1	FDX1_HABIN		P44451 haemophilus
5	41	45.1	434	1	SYN_PYRAB		Q9v251 pyrococcus
6	40	44.0	260	1	AG22_SHEEP		Q28929 ovis aries
7	40	44.0	351	1	SGF3_BOMMO		Q17237 bombyx mori
8	40	44.0	540	1	RIK2_HUMAN		Q43353 h receptor
9	40	44.0	2517	1	NCR2_HUMAN		Q9v618 h nuclear r
10	39	42.9	85	1	HEPC_MORCS		P82951 morone chry
11	39	42.9	351	1	DESA_SYNY3		P20388 synechocyst
12	39	42.9	508	1	GALT_BACHD		Q8kdv2 bacillus ha
13	39	42.9	539	1	RIK2_MOUSE		P58801 mus musculus
14	39	42.9	566	1	MASY_CUCMA		P24571 cucurbita m
15	39	42.9	568	1	MASY_CUCSA		P08216 cucumis sat
16	39	42.9	1004	1	CARE_HUMAN		Q9bx16 homo sapien
17	39	42.9	1021	1	Y2R2_DROME		P16425 drosophila
18	39	42.9	1312	1	DPOL_PYRSD		Q51334 pyrococcus
19	38	41.8	64	1	RL35_ECOLI		P07085 escherichia
20	38	41.8	274	1	MEPA_ECOLI		P14007 escherichia
21	38	41.8	358	1	RECF_PASMU		Q9clg6 pasteurella
22	38	41.8	420	1	KC2A_THEPA		P28547 theileria p
23	38	41.8	427	1	CG23_YEAST		P24870 saccharomyc
24	38	41.8	434	1	SYN_PYRFU		Q8u4d3 pyrococcus
25	38	41.8	434	1	SYN_PYRHO		Q57980 pyrococcus
26	38	41.8	459	1	RSP6_CHLRE		Q01657 chlamydomon
27	38	41.8	488	1	SUOX_HUMAN		P51687 homo sapien
28	38	41.8	881	1	HELI_HSVEB		P28934 equine herp
29	38	41.8	881	1	HELI_V2VD		P09303 varicella-z
30	38	41.8	943	1	ODOL_AZOVI		P20707 azotobacter
31	38	41.8	2054	1	YCF2_PINTH		P41653 pinus thunb
32	38	41.8	3430	1	POLG_WNV		P06935 w genome po
33	38	41.8	3433	1	POLG_KUNJM		P14335 k genome po

#### RESULT 1

ID	CCT1_MOUSE	STANDARD	PRT	724 AA
AC	Q9QWV9; Q9ZOU7;			
DC	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Cyclin T1 (Cyclin T) (CyclT).			
GN	CCNT1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALE/C;			
RC	MEDLINE=99059742; Pubmed=9843510;			
RA	Bieniasz P.D., Grdina T.A., Bogerd H.P., Cullen B.R.;			
RT	"Recruitment of a protein complex containing Tat and cyclin T1 to TAR			
RT	governs the species specificity of HIV-1 Tat.;"			
RL	EMBO J. 17:7056-7065(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CS7BL/6; TISSUE=Brain;			
RC	MEDLINE=99051315; Pubmed=9832504;			
RA	Garber M.E., Wei P., Kewairamani V.N., Mayall T.P., Herrmann C.H.,			
RA	Rice A.P., Littman D.R., Jones K.A.;			
RT	"The interaction between HIV-1 Tat and human cyclin T1 requires zinc			
RT	and a critical cysteine residue that is not conserved in the murine			
RT	CyclT1 protein.;"			
RL	Genes Dev. 12:3512-3527(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS.			
RC	TISSUE=Spleen;			
RC	MEDLINE=99263519; Pubmed=10329126;			
RA	Kwak Y.T., Ivanov D., Guo J., Nee E., Gaynor R.B.;			
RT	"Role of the human and murine cyclin T proteins in regulating HIV-1			
RT	Tat-activation.;"			
RL	J. Mol. Biol. 288:57-69(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS.			
RC	TISSUE=Fibroblast;			
RC	MEDLINE=99145542; Pubmed=9990016;			
RA	Fujinaga K., Taube R., Wimmer J., Cujec T.P., Peterlin B.M.;			
RT	"Interactions between human cyclin T, Tat, and the transactivation			
RT	response element (TAR) are disrupted by a cysteine to tyrosine			
RT	substitution found in mouse cyclin T.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:1285-1290(1999).			
CC	-1- FUNCTION: REGULATORY SUBUNIT OF THE CYCLIN-DEPENDENT KINASE PAIR			
CC	(CDK9/CYCLIN T) COMPLEX, ALSO CALLED POSITIVE TRANSCRIPTION			
CC	ELONGATION FACTOR B (P-TEFb), WHICH IS PROPOSED TO FACILITATE THE			
CC	TRANSITION FROM ABORTIVE TO PRODUCTION ELONGATION BY			
CC	PHOSPHORYLATING THE CTD (CARBOXY-TERMINAL DOMAIN) OF THE LARGE			
CC	SUBUNIT OF RNA POLYMERASE II (RNAP II). DOES NOT BIND EFFICIENTLY			
CC	TO THE TRANSCRIPTIONAL DOMAIN OF THE HIV-1 NUCLEAR TRANSCRIPTIONAL			
CC	ACTIVATOR, TAT.			

#### ALIGNMENTS

34	37	40.7	102	1	YACL_MAIZE	P08771 zea mays (m
35	37	40.7	112	1	YE14_YEAST	P39973 saccharomyc
36	37	40.7	207	1	NU6M_PROMI	Q37626 saccharomyc
37	37	40.7	273	1	TC1A_CABEL	P03934 caenorhabdi
38	37	40.7	365	1	CMLE_NEUCR	P38677 neurospora
39	37	40.7	395	1	SYT8_MOUSE	Q9r0n6 mus musculu
40	37	40.7	411	1	LAG1_YEAST	P38703 saccharomyc
41	37	40.7	426	1	PYRC_THEAQ	P36081 thermus aqu
42	37	40.7	498	1	GSHB_SCHPO	P35669 schizosacch
43	37	40.7	529	1	YB98_HUMAN	Q9ulm2 homo sapien
44	37	40.7	632	1	YO22_CABEL	P34672 caenorhabdi
45	37	40.7	653	1	PPE1_HUMAN	O14829 homo sapien

CC -|- SUBUNIT: ASSOCIATES WITH CDK9 TO FORM P-TEFB. CYCLIN T1 IS THE  
 CC -|- PREDOMINANT CYCLIN ASSOCIATED WITH CDK9.  
 CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -|- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. STRONG TO CYCLIN C  
 CC SUBFAMILY.

CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----

DR EMBL: AF095640; AAD13656.1; -  
 DR EMBL: AF109179; AAD19654.1; -  
 DR EMBL: AF087662; AAD17798.1; -  
 DR EMBL: AF113951; AAD17205.1; -  
 DR MGD: MGI:1328363; Ccnt1.  
 DR InterPro: IPR004366; Ccnt1.  
 DR Pfam: PF00134; cyclin; 1.  
 DR SMART: SM00385; CYCLIN; 1.  
 DR PROSITE: PS00292; CYCLINS; FALSE\_NEG.  
 KW Cyclin; Cell cycle; Cell division; Coiled coil;  
 KW Transcription regulation; Nuclear protein.  
 FT DOMAIN 253 270 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 254 272 TAT-TAR RECOGNITION MOTIF (TRM).  
 FT DOMAIN 319 322 POLY-SER.  
 FT DOMAIN 384 425 COILED COIL (POTENTIAL).  
 FT DOMAIN 516 525 POLY-HIS.  
 FT DOMAIN 565 569 POLY-SER.  
 FT DOMAIN 715 723 POLY-PRO.  
 FT CONFLICT 48 48 V -> M (IN REF. 3).  
 SQ SEQUENCE 724 AA; 80565 MW; 7396E428F5A8B91B CRC64;

Query Match 49.5%; Score 45; DB 1; Length 724;  
 Best Local Similarity 58.3%; Pred. No. 6.9;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 YFSHRHAHLRSP 15  
 DB 522 HHSHRHSLQLP 533

RESULT 2  
 TNF5\_CHICK STANDARD; PRT; 272 AA.  
 AC Q918D8;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CDA40 ligand) (CDA40-  
 DE L) (CD154 protein).  
 GN TNFSF5 OR CD40LG OR CD40L.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-White leghorn; TISSUE=Spleen;  
 RA Tregaskes C.A., Young J.R., Burnside J.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -|- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell  
 CC proliferation in the absence of co-stimulus as well as IgE  
 CC production in the presence of IL-4. Involved in immunoglobulin  
 CC class switching (By similarity).  
 CC -|- SUBUNIT: Homotrimer (By similarity).  
 CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an  
 CC extracellular soluble form (By similarity).  
 CC -|- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (By similarity).

CC -|- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----

DR EMBL: AJ243435; CAB95748.1; -  
 DR HSSP: P29965; ITALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 272 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY.  
 FT CHAIN 111 272 MEMBER 5, MEMBRANE FORM.  
 FT DOMAIN 1 23 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT TRANSMEM 24 44 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 45 272 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 FT SITE 110 111 EXTRACELLULAR (POTENTIAL).  
 FT DISULFID 190 229 CLEAVAGE (BY SIMILARITY).  
 FT CARBOHYD 124 124 POTENTIAL.  
 FT CARBOHYD 146 146 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 272 AA; 30862 MW; 5409F24A8E53CCD7 CRC64;

Query Match 46.2%; Score 42; DB 1; Length 272;  
 Best Local Similarity 50.0%; Pred. No. 7.5;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRS 14  
 DB 109 FEMHRSHEPHLKS 122

RESULT 3  
 TRX2\_ECOLI STANDARD; PRT; 248 AA.  
 ID TRX2\_ECOLI  
 AC P22710;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Trax protein.  
 GN TRAX.  
 OS Escherichia coli.  
 OG Plasmid IncFII R100, and Plasmid R6-5.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID-INC FII R100;  
 RX MEDLINE=90317835; PubMed=2164585;  
 RA Yoshioka Y., Fujita Y., Ohtsubo E.;  
 RT "Nucleotide sequence of the promoter-distal region of the tra operon  
 RL of plasmid R100, including tral (DNA helicase I) and traf genes.";  
 RL J. Mol. Biol. 214:39-53(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID-R6-5;  
 RX MEDLINE=92009201; PubMed=1916281;  
 RA Cram D.S., Loh S.M., Cheah K.C., Skurray R.A.;



DR PRINTS; PRO1042; TRNASYNTHASP.  
 DR TIGRFAMS; TIGR00457; asns; 1.  
 DR PROSITE; PS50862; AA\_TRNA\_LIGASE\_II; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 434 AA; 50242 MW; 9A08BBCE2458A6A3 CRC64;  
 Query Match 45.1%; Score 41; DB 1; Length 434;  
 Best Local Similarity 46.2%; Pred. No. 18;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 MYFSHRHAHLRSP 15  
 : :||| :|||  
 Db 117 LLLDYRHHLRSP 129

RESULT 6  
 AG22\_SHEEP STANDARD; PRT; 260 AA.  
 ID AG22\_SHEEP  
 AC Q28929;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Type-2 angiotensin II receptor (AT2) (Fragment).  
 GN AGTR2.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Kidney;  
 RX MEDLINE=96169644; PubMed=8618791;  
 RA Robillard J.E., Page W.V., Mathews M.S., Schutte B.C., Nuyt A.M.,  
 RA Segar J.L.;  
 RT "Differential gene expression and regulation of renal angiotensin II  
 receptor subtypes (AT1 and AT2) during fetal life in sheep.";  
 RL Pediatr. Res. 38:896-904(1995).  
 CC -1- FUNCTION: RECEPTOR FOR ANGIOTENSIN II. MAY HAVE A ROLE IN CELL  
 MORPHOGENESIS AND RELATED EVENTS IN GROWTH AND DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; S81979; AAB36404.1; -  
 DR HSSP; P34996; 1DDO.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 NON\_TER 1 24  
 FT TRANSMEM <1 24 1 (POTENTIAL).  
 FT DOMAIN 25 33 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 34 55 2 (POTENTIAL).  
 FT DOMAIN 56 72 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 73 93 3 (POTENTIAL).  
 FT DOMAIN 94 113 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 114 132 4 (POTENTIAL).  
 FT DOMAIN 133 161 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 162 187 5 (POTENTIAL).  
 FT DOMAIN 188 209 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 210 231 6 (POTENTIAL).  
 FT DOMAIN 232 238 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 239 >260 7 (POTENTIAL).

FT NON\_TER 260 260  
 SQ SEQUENCE 260 AA; 29549 MW; BB5B48FF029A0B42 CRC64;  
 Query Match 44.0%; Score 40; DB 1; Length 260;  
 Best Local Similarity 40.0%; Pred. No. 15;  
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
 : :||| :|||  
 Db 53 WATYYSRHYDWIRGP 67

RESULT 7  
 SGF3\_BOMMO STANDARD; PRT; 351 AA.  
 ID SGF3\_BOMMO  
 AC Q17237;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Silk gland factor 3 (SGF-3) (POU domain protein M1).  
 GN SGF3 OR POU-M1  
 OS Bombyx mori (Silk moth).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;  
 OC Ditrysia; Bombycoidea; Bombycidae; Bombyx.  
 OX NCBI\_TaxID=7091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Kinshu X Showa; TISSUE=Silk gland;  
 RX MEDLINE=93374935; PubMed=7690034;  
 RA Fukuta M., Matsuno K., Hui C.-C., Nagata T., Takiya S., Xu P.-X.,  
 RA Ueno K., Suzuki Y.;  
 RT "Molecular cloning of a POU domain-containing factor involved in the  
 regulation of the Bombyx sericin-1 gene.";  
 RL J. Biol. Chem. 268:19471-19475(1993).  
 CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF SERICIN-1  
 GENE.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE MIDDLE SILK GLAND.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION DIFFERENTIALLY REGULATED IN THE  
 POSTERIOR AND MIDDLE SILK GLANDS DURING THE FOURTH MOLT/FIFTH  
 INTERMOLT.  
 CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.  
 CC CLASS-3 SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M64781; AAA27841.1; -  
 DR HSSP; P14859; 1OCT.  
 DR TRANSFAC; T00746; -  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000327; POU\_domain.  
 DR Pfam; PF00046; homeobox; 1.  
 DR Pfam; PF00157; pou; 1.  
 DR PRINTS; PR00028; POUDOMAIN.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR ProDom; PD000583; POU\_domain; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR SMART; SM00352; POU; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 DR PROSITE; PS00035; POU\_1; 1.  
 DR PROSITE; PS00465; POU\_2; 1.  
 KW Homeobox; DNA-binding; Transcription regulation; Activator;  
 KW Nuclear protein; Neurogenesis.  
 FT DOMAIN 152 223 POU.



```
FT DNA_BIND 241 300 HOMEBOX.
SQ SEQUENCE 351 AA; 38591 MW; A6DFA4A59E9322B6 CRC64;

Query Match 44.0%; Score 40; DB 1; Length 351;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WRMYFSHRAH 11
Db 64 WAMHQHAAH 74

RESULT 8
ID RIK2_HUMAN STANDARD; PRT; 540 AA.
AC 043353;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Receptor-interacting serine/threonine protein kinase 2 (EC 2.7.1.-)
DE (RIP-2) (CARD-containing interleukin-1 beta converting enzyme
DE associated kinase) (CARD-containing IL-1 beta ICE-kinase).
GN RIPK2 OR RICK OR RIP2 OR CARDIAK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID-9606;
RN [1]

SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-47.
RX MEDLINE-98241596; PubMed-9575181;
RA Inohara N., del Peso L., Koseki T., Chen S., Nunez G.;
RT "RICK, a novel protein kinase containing a caspase recruitment domain,
RT interacts with CLARP and regulates CD95-mediated apoptosis.";
RL J. Biol. Chem. 273:12296-12300(1998).

[2]
SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-47.
RN
RP
RC TISSUE-Endothelial cells;
RX MEDLINE-98307936; PubMed-9642260;
RA McCarthy J.V., Ni J., Dixit V.M.;
RT "RIP2 is a novel NF-kappaB-activating and cell death-inducing
RT kinase.";
RL J. Biol. Chem. 273:16968-16975(1998).

[3]
SEQUENCE FROM N.A., AND MUTAGENESIS OF ASP-146.
RN
RP
RX MEDLINE-98381580; PubMed-9705938;
RA Thome M., Hofmann K., Burns K., Martinon F., Bodmer J.-L.,
RA Matmann C., Tschoep J.;
RT "Identification of CARDIAK, a RIP-like kinase that associates with
RT caspase-1.";
RL Curr. Biol. 8:885-888(1998).

[4]
SEQUENCE FROM N.A.
RN
RP
RA Ozersky P., Holmes A., Broy M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

[5]
SEQUENCE FROM N.A.
RN
RP
RA Platzner M., Varon R.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

[6]
SEQUENCE FROM N.A.
RN
RP
RC TISSUE-Skin;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Activates pro-caspase-1 and pro-caspase-8. Potentiates
CC CASP-8-mediated apoptosis. Activates NF-kappaB.
CC -1- SUBUNIT: Binds to CFLAR/CLARP and CASP1 via their CARD domains.
CC Binds to BIRC3/c-IAP1 and BIRC2/c-IAP2, TRAF1, TRAF2, TRAF5 and
CC TRAF6. May be a component of both the TNFRSF1A and TNFRSF5/CD40
CC receptor complex.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: Detected in heart, brain, placenta, lung,
CC peripheral blood leukocytes, spleen, kidney, testis, prostate,
CC
```

```
CC pancreas and lymph node.
CC -1- PTM: Autophosphorylated.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 CARD DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF027706; AAC34970.1; -
DR EMBL; AF078530; AAC27722.1; -
DR EMBL; AF064824; AAC25668.1; -
DR EMBL; AC004003; AAC24561.1; -
DR EMBL; AF117829; AAD04634.1; -
DR EMBL; BC004553; AAH04553.1; -
DR Genew; HGNC:10020; RIPK2.
DR MIM; 603455; -
DR InterPro; IPR001315; CARD.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR004040; STY_kinase.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00114; CARD; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS50209; CARD; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Apoptosis.
FT DOMAIN 18 294 PROTEIN KINASE.
FT DOMAIN 432 524 CARD.
FT NP_BIND 24 32 ATP (BY SIMILARITY).
FT BINDING 47 47 ATP.
FT ACT_SITE 146 146
FT MUTAGEN 47 47 K->A: ABOLISHES KINASE ACTIVITY.
FT MUTAGEN 47 47 K->M: REDUCES FAS-MEDIATED APOPTOSIS.
FT MUTAGEN 146 146 D->N: ABOLISHES KINASE ACTIVITY.
SQ SEQUENCE 540 AA; 61194 MW; 575A692239505792 CRC64;

Query Match 44.0%; Score 40; DB 1; Length 540;
Best Local Similarity 33.3%; Pred. No. 33;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRAHLRSP 15
Db 40 WRVQVAVKHLHTTP 54

RESULT 9
NCR2_HUMAN
ID NCR2_HUMAN STANDARD; PRT; 2517 AA.
AC Q9Y618; Q9Y5U0; Q13354; O00613; O15416;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nuclear receptor co-repressor 2 (N-COR2) (Silencing mediator of
DE retinoic acid and thyroid hormone receptor) (SMRT) (Thyroid-,
DE retinoic-acid-receptor-associated co-repressor) (TRAC) (CTG26).
DE associating factor) (TRAC) (CTG26).
GN NCR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SMRT).
RC TISSUE-Pituitary;
```

RX MEDLINE-99178941; PubMed-10077563;  
 RA Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M.;  
 RT "Unique forms of human and mouse nuclear receptor corepressor SMRT.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2639-2644(1999).  
 RN [2].  
 RP SEQUENCE FROM N.A. (ISOFORM SMRT).  
 RC TISSUE=Cervical adenocarcinoma;  
 RX MEDLINE-99199215; PubMed-10097068;  
 RA Park E.J., Schroen D.J., Yang M., Li H., Li L., Chen J.D.;  
 RT "SMRte, a silencing mediator for retinoid and thyroid hormone  
 receptors-extended isoform that is more related to the nuclear  
 receptor corepressor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:3519-3524(1999).  
 RN [3].  
 RP SEQUENCE OF 1023-2517 FROM N.A.  
 RC TISSUE=Cervical adenocarcinoma;  
 RX MEDLINE-96008552; PubMed-7566127;  
 RA Chen J.D., Evans R.M.;  
 RT "A transcriptional co-repressor that interacts with nuclear hormone  
 receptors.";  
 RL Nature 377:454-457(1995).  
 RN [4].  
 RP SEQUENCE FROM N.A. (ISOFORM TRAC-1).  
 RC TISSUE=Fetal liver;  
 RX MEDLINE-96408715; PubMed-8813722;  
 RA Sande S., Privalsky M.L.;  
 RT "Identification of TRACs (T3 receptor-associating cofactors), a family  
 of cofactors that associate with, and modulate the activity of,  
 nuclear hormone receptors.";  
 RL Mol. Endocrinol. 10:813-825(1996).  
 RN [5].  
 RP SEQUENCE OF 428-613 FROM N.A.  
 RC TISSUE=Brain cortex;  
 RX MEDLINE-97369492; PubMed-9225980;  
 RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.H., Kidwai A.S.,  
 RA Breschel T.S., Stine O.C., Callahan C., McInnis M.G., Ross C.A.;  
 RT "cDNAs with long CAG trinucleotide repeats from human brain.";  
 RL Hum. Genet. 100:114-122(1997).  
 CC -1- FUNCTION: MEDIATES THE TRANSCRIPTIONAL REPRESSION ACTIVITY OF SOME  
 CC NUCLEAR RECEPTORS BY PROMOTING CHROMATIN CONDENSATION, THUS  
 CC PREVENTING ACCESS OF THE BASAL TRANSCRIPTION.  
 CC -1- SUBUNIT: FORMS A LARGE COREPRESSOR COMPLEX THAT CONTAINS SIN3A/B  
 CC AND HISTONE DEACETYLASES HDAC1 AND HDAC2. THIS COMPLEX ASSOCIATES  
 CC WITH THE THYROID (TR) AND THE RETINOID ACID RECEPTORS (RAR) IN THE  
 CC ABSENCE OF LIGAND, AND MAY STABILIZE THEIR INTERACTION WITH TFIIIB.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SMRT/TRAC-2 (SHOWN HERE) AND  
 CC TRAC-1; ARE PRODUCED BY ALTERNATIVE SPLICING. TRAC-1 CONTAINS ONLY  
 CC THE C-TERMINAL RECEPTOR-INTERACTING DOMAIN AND ACTS AS AN  
 CC ANTIREPRESSOR.  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS. HIGH LEVELS OF EXPRESSION ARE  
 CC DETECTED IN LUNG, SPLEEN AND BRAIN.  
 CC -1- INDUCTION: REGULATED DURING CELL CYCLE PROGRESSION.  
 CC -1- DOMAIN: THE N-TERMINAL REGION CONTAINS REPRESSION FUNCTIONS THAT  
 CC ARE DIVIDED INTO THREE INDEPENDANT REPRESSION DOMAINS (RD1, RD2  
 CC AND RD3). THE C-TERMINAL REGION CONTAINS THE NUCLEAR RECEPTOR-  
 CC INTERACTING DOMAINS THAT ARE DIVIDED IN TWO SEPARATE INTERACTION  
 CC DOMAINS (ID1 AND ID2).  
 CC -1- DOMAIN: THE TWO INTERACTION DOMAINS (ID) CONTAIN A CONSERVED  
 CC SEQUENCE REFERRED TO AS THE CORNR BOX. THIS MOTIF IS REQUIRED AND  
 CC SUFFICIENT TO PERMIT BINDING TO UNLIGANDED TR AND RARS. SEQUENCES  
 CC FLANKING THE CORNR BOX DETERMINE NUCLEAR HORMONE RECEPTOR  
 CC SPECIFICITY.  
 CC -1- SIMILARITY: CONTAINS 1 SANT-A DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 CORNR BOXES.  
 CC -1- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF113003; AAD20946.1; -;  
 DR EMBL: AF125672; AAD22973.1; -;  
 DR EMBL: U37146; AAC50236.1; -;  
 DR EMBL: S83390; AAB50847.1; -;  
 DR EMBL: U80750; AAB91446.1; -;  
 DR TRANSFAC: T04689; -;  
 DR Genew: HGNC:7673; NCOR2.  
 DR MIM: 600848; -;  
 DR InterPro: IPR001005; Myb\_DNA\_binding.  
 DR Pfam: PF00249; myb\_DNA-binding; 2.  
 DR SMART: SM00395; SANT; 2.  
 DR PROSITE: PS50090; MYB\_3; 1.  
 KW Nuclear protein; Transcription regulation; DNA-binding; Repressor;  
 KW Coiled coil; Alternative splicing.  
 FT DOMAIN 174 215 COILED COIL (POTENTIAL).  
 FT DOMAIN 254 312 INTERACTION WITH SIN3A/B (BY SIMILARITY).  
 FT DNA\_BIND 429 474 MYB.  
 FT DOMAIN 522 561 COILED COIL (POTENTIAL).  
 FT DOMAIN 778 820 PRO-RICH.  
 FT DOMAIN 2139 2143 CORNR BOX OF ID1.  
 FT DOMAIN 2342 2346 CORNR BOX OF ID2.  
 FT DOMAIN 494 510 POLY-GLN.  
 FT DOMAIN 682 685 POLY-LVS.  
 FT DOMAIN 994 1002 POLY-PRO.  
 FT DOMAIN 1384 1389 POLY-PRO.  
 FT DOMAIN 1842 1846 POLY-GLY.  
 FT DOMAIN 2479 2482 POLY-PRO.  
 FT VARSPLIC 1 1702 MISSING (IN ISOFORM TRAC-1).  
 FT CONFLICT 2353 2398 MISSING (IN ISOFORM TRAC-1).  
 FT CONFLICT 7 7 L -> P (IN REF. 2).  
 FT CONFLICT 295 295 K -> E (IN REF. 2).  
 FT CONFLICT 309 309 L -> W (IN REF. 2).  
 FT CONFLICT 352 352 MISSING (IN REF. 2).  
 FT CONFLICT 365 365 A -> P (IN REF. 2).  
 FT CONFLICT 612 613 SS -> EF (IN REF. 5).  
 FT CONFLICT 711 711 S -> T (IN REF. 2).  
 FT CONFLICT 724 740 MISSING (IN REF. 2).  
 FT CONFLICT 787 796 RRTSRAPIEP -> PEDIPAPTES (IN REF. 2).  
 FT CONFLICT 804 804 G -> L (IN REF. 2).  
 FT CONFLICT 814 814 S -> F (IN REF. 2).  
 FT CONFLICT 817 817 A -> S (IN REF. 2).  
 FT CONFLICT 889 889 G -> R (IN REF. 2).  
 FT CONFLICT 1023 1030 SRSPAPPA -> MEAWDAHP (IN REF. 3).  
 FT CONFLICT 1034 1034 A -> AEKPVFFPA (IN REF. 2).  
 FT CONFLICT 1894 1894 K -> T (IN REF. 4).  
 FT CONFLICT 2494 2494 P -> A (IN REF. 4).  
 SQ SEQUENCE 2517 AA; 274031 MW; F5805C01761258C0 CRC64;  
 Query Match 44.0%; Score 40; DB 1; Length 2517;  
 Best Local Similarity 70.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 6 SHRAHLRSP 15  
 DB 1853 SHSHAHQHP 1862  
 RESULT 10  
 ID HEPC\_MORCS STANDARD; PRT; 85 AA.  
 AC P82951;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hepcidin precursor.  
 OS Morone chrysops x Morone saxatilis (White bass x Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
OX Moronidae; Morone.  
RN NCBI\_TaxID=45352;  
RP SEQUENCE FROM N.A., SEQUENCE OF 65-85, TISSUE SPECIFICITY, AND  
RP ANTIOTIC ACTIVITY.  
RC TISSUE=GI, and skin;  
RX MEDLINE=21982021; PubMed=11985602;  
RA Shike H., Lauth X., Westerman M.E., Ostland V.E., Carlberg J.M.,  
RA Van Oost J.C., Shimizu C., Bulet P., Burns J.C.;  
RT "Bass hepcidin is a novel antimicrobial peptide induced by bacterial  
RT challenge";  
RL Eur. J. Biochem. 269:2232-2237(2002).  
CC -!- FUNCTION: Seems to act as a signaling molecule involved in the  
CC maintenance of iron homeostasis. Seems to be required in  
CC conjunction with HFE to regulate both intestinal iron absorption  
CC and iron storage in macrophages (By similarity).  
CC -!- FUNCTION: Antimicrobial activity against Gram-negative bacteria  
CC such as E.coli.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Predominantly expressed in liver.  
CC -!- MASS SPECTROMETRY: MW=2255.97; METHOD=MALDI; RANGE=65-85.  
CC -!- INDUCTION: By bacterial challenge.  
CC -!- SIMILARITY: BELONGS TO THE HEPICIDIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AF394246; AAM28440.1; -  
KW Antibiotic; Hormone; Signal.  
FT SIGNAL 1 24  
FT PROPEP 25 64  
FT CHAIN 65 85  
FT DISULFD 66 72  
FT DISULFD 69 82  
FT DISULFD 70 83  
FT DISULFD 73 79  
SQ SEQUENCE 85 AA; 9484 MW; 0FEA55CF0A522C84 CRC64;  
Query Match 42.9%; Score 39; DB 1; Length 85;  
Best Local Similarity 40.0%; Pred. No. 6.6;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 WRMYFSHRHAHLRSP 15  
Db 49 WKPNYNRRHKRHSSP 63  
RESULT 11  
ID DESA\_SYNY3 STANDARD; PRT; 351 AA.  
AC P20388;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fatty acid desaturase (EC 1.14.99.-) (Delta 12 desaturase).  
DE DESA OR SLR1350.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90370121; PubMed=2118597;  
RA Wada H., Gombos Z., Murata N.;  
RT "Enhancement of chilling tolerance of a cyanobacterium by genetic  
RT manipulation of fatty acid desaturation.";  
RL Nature 347:200-203(1990).  
RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,  
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
RA Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
CC -!- FUNCTION: CAN INTRODUCE A SECOND CIS DOUBLE BOND AT THE DELTA 12  
CC POSITION OF FATTY ACID BOUND TO MEMBRANE GLYCEROLIPIDS. THIS  
CC ENZYME IS INVOLVED IN CHILLING TOLERANCE BECAUSE THE PHASE  
CC TRANSITION TEMPERATURE OF LIPIDS OF CELLULAR MEMBRANES DEPENDS ON  
CC THE DEGREE OF UNSATURATION OF FATTY ACIDS OF THE MEMBRANE LIPIDS.  
CC -!- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X53508; CAA37584.1; -  
DR EMBL; D90912; BAA18169.1; -  
PIR; S11519; S11519.  
DR InterPro; IPR001225; FA\_desaturase.  
DR Pfam; PF00487; FA\_desaturase; 1.  
DR ProDom; PD001081; FA\_desaturase; 1.  
DR PROSITE; PS00574; FATTY-ACID-DESATUR\_2; 1.  
KW Oxidoreductase; Fatty acid biosynthesis; Membrane; Complete proteome.  
FT DOMAIN 90 94  
FT DOMAIN 126 130  
FT DOMAIN 287 291  
SQ SEQUENCE 351 AA; 40495 MW; 85DBEE341F73F77 CRC64;  
Query Match 42.9%; Score 39; DB 1; Length 351;  
Best Local Similarity 45.5%; Pred. No. 30;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 WRMYFSHRHAH 11  
Db 122 WRLLDHHHLH 132  
RESULT 12  
ID GALT\_BACHD STANDARD; PRT; 508 AA.  
AC Q9KDV2; Q9RC74;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10) (Gal-1-P  
DE uridylyltransferase).  
GN GALT OR BHL109.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki K., Masui N.,  
RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
RN [2]  
RP SEQUENCE OF 1-450 FROM N.A.

```
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=99411980; PubMed=10484179;
RA Takami H., Takaki Y., Nakasone K., Sakiyama T., Maeno G., Sasaki R.,
RA Hirama C., Fuji F., Masui N.;
RT Genetic analysis of the chromosome of alkaliphilic Bacillus
RL halodurans C-125.;
CC -|- CAPALYTIC ACTIVITY: UTP + alpha-D-galactose 1-phosphate =
CC diophosphate + UDP-galactose.
CC -|- PATHWAY: Galactose metabolism; second step.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -|- SIMILARITY: BELONGS TO THE GALACTOSE-1-PHOSPHATE
CC URIDYLTRANSFERASE FAMILY 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AP001510; BAB04828.1; -.
CC EMBL: AB024554; BAA83925.1; -.
CC InterPro: IPR000880; GalP_UDP_transf.
CC Pfam: PF01087; GalP_UDP_transf_II.
CC Pfam: PF02744; GalP_UDP_tr.C; 1.
CC TIGRFAMs: TIGR01239; galT_2; 1.
CC PROSITE: PS01163; GAL_P_UDP_TRANSF_II; 1.
CC Transferrase; Nucleotidyltransferase; Galactose metabolism;
CC Complete proteome.
CC SEQUENCE 508 AA; 57989 MW; 16AF6F607FCEAE2E CRC64;
CC
CC Query Match 42.9%; Score 39; DB 1; Length 508;
CC Best Local Similarity 50.0%; Pred. No. 45;
CC Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
CC
CC QY 1 WRMYFSHRHAHL 12
CC I : : : : :
CC DB 451 WALAIKERHAHL 462
CC
CC RESULT 13
CC RIK2_MOUSE STANDARD; PRT; 539 AA.
CC AC P58801;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Receptor-interacting serine/threonine protein kinase 2 (EC 2.7.1.-).
CC GN RIKP2.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=C57BL/6;
CC MEDLINE=21891093; PubMed=11894097;
CC RA Chin A.I., Dempsey P.W., Bruhn K., Miller J.F., Xu Y., Cheng G.;
CC RT Involvement of receptor-interacting protein 2 in innate and adaptive
CC immune responses."
CC RL Nature 416:190-194(2002).
CC -|- FUNCTION: Activates pro-caspase-1 and pro-caspase-8. Potentiates
CC CASP-8-mediated apoptosis. Activates NF-kappaB (By similarity).
CC -|- SUBUNIT: Binds to CFLAR/CLARP and CASP1 via their CARD domains.
CC Binds to BIRC2/c-IAP1 and BIRC2/c-IAP2, TRAF1, TRAF2, TRAF5 and
CC TRAF6. May be a component of both the TNFRSF1A and TNFRSF5/CD40
CC receptor complex (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- PTM: Autophosphorylated (By similarity).
CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
```

```
CC -|- SIMILARITY: CONTAINS 1 CARD DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF461040; AAL96436.1; -.
CC PROSITE: PS50209; CARD; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation; Apoptosis.
CC FT DOMAIN 18 294 PROTEIN KINASE.
CC FT DOMAIN 431 523 CARD.
CC FT NP_BIND 24 32 ATP (BY SIMILARITY).
CC FT BINDING 47 47 .ATP (BY SIMILARITY).
CC FT ACT_SITE 146 146 BY SIMILARITY.
CC SQ SEQUENCE 539 AA; 60400 MW; 42951BF97CA15DFA CRC64;
CC
CC Query Match 42.9%; Score 39; DB 1; Length 539;
CC Best Local Similarity 33.3%; Pred. No. 48;
CC Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
CC
CC QY 1 WRMYFSHRHAHLRSP 15
CC I : : : : :
CC DB 40 WRVRVAVKHLHIHTP 54
CC
CC RESULT 14
CC MASV_CUCMA STANDARD; PRT; 566 AA.
CC ID P24571;
CC DT 01-MAR-1992 (Rel. 21, Created)
CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE Malate synthase, glyoxysomal (EC 4.1.3.2).
CC OS Cucurbita maxima (Pumpkin) (Winter squash).
CC OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
CC OX NCBI_TaxID=3661;
CC [1]
CC SEQUENCE FROM N.A., AND SEQUENCE OF 83-91: 197-206 AND 417-426.
CC STRAIN=cv. Kurokawa Anakuri Nankin; TISSUE=Etioiated cotyledon;
CC RX MEDLINE=91224123; PubMed=1709098;
CC RA Mori H., Takeda-Yoshikawa Y., Hara-Nishimura I., Nishimura M.;
CC RT "Pumpkin malate synthase. Cloning and sequencing of the cDNA and
CC northern blot analysis."
CC RL Eur. J. Biochem. 197:331-336(1991).
CC -|- CATALYTIC ACTIVITY: L-malate + CoA -> acetyl-CoA + H(2)O +
CC glyoxylate.
CC -|- PATHWAY: SECOND STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
CC TRICARBOXYLIC ACID CYCLE (IN BACTERIA, FUNGI AND PLANTS).
CC -|- SUBCELLULAR LOCATION: Glyoxysomal.
CC -|- SIMILARITY: BELONGS TO THE MALATE SYNTHASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X56948; CAA40262.1; -.
CC PIR: S15387; S15387.
CC PIR: S14601; S14601.
CC InterPro: IPR001465; Malate_synthase.
```

DR Pfam: PF01274; Malate synthase; 1.  
 DR PROSITE: PS00342; MICROBODIES\_CTER; 1.  
 DR PROSITE: PS00510; MALATE\_SYNTHASE; 1.  
 KW Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.  
 FT SITE 564 566 MICROBODY TARGETING SIGNAL (POTENTIAL).  
 SQ SEQUENCE 566 AA: 64635 MW: 877777DID593834A CRC64;

Query Match 42.9%; Score 39; DB 1; Length 566;  
 Best Local Similarity 54.5%; Pred. No. 50;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 3 MYFSHRHAHLR 13  
 :||| ||| :  
 Db 209 LYFFHNHANFR 219

RESULT 15  
 MASY\_CUCSA STANDARD: PRT; 568 AA.  
 AC P08216;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Malate synthase, glyoxysomal (EC 4.1.3.2).  
 OS Cucumis sativus (Cucumber).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=3659;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. LONG GREEN RIDGE;  
 RX MEDLINE=91370851; PubMed=2491683;  
 RA Graham I.A., Smith L.M., Brown J.W.S., Leaver C.J., Smith S.M.;  
 RT "The malate synthase gene of cucumber.";  
 RL Plant Mol. Biol. 13:673-684(1989).  
 RN [2]  
 RP SEQUENCE OF 475-568 FROM N.A.  
 RA Smith S.M., Leaver C.J.;  
 RT "Glyoxysomal malate synthase of cucumber: molecular cloning of a cDNA  
 and regulation of enzyme synthesis during germination.";  
 RL Plant Physiol. 81:762-767(1986).  
 CC -|- CATALYTIC ACTIVITY: L-malate + CoA = acetyl-CoA + H(2)O +  
 glyoxylate.  
 CC -|- PATHWAY: SECOND STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE  
 TRICARBOXYLIC ACID CYCLE (IN BACTERIA, FUNGI AND PLANTS).  
 CC -|- SUBCELLULAR LOCATION: Glyoxysomal.  
 CC -|- SIMILARITY: BELONGS TO THE MALATE SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X15425; CAA33465.1; -;  
 DR EMBL: M16219; AAA33123.1; -;  
 DR PIR: S07550; SYKVMA.  
 DR InterPro: IPR001465; Malate\_synthase.  
 DR Pfam: PF01274; Malate\_synthase; 1.  
 DR PROSITE: PS00342; MICROBODIES\_CTER; 1.  
 DR PROSITE: PS00510; MALATE\_SYNTHASE; 1.  
 KW Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.  
 FT SITE 566 568 MICROBODY TARGETING SIGNAL (POTENTIAL).  
 FT CONFLICT 475 475 I -> G (IN REF. 2).  
 SQ SEQUENCE 568 AA: 64961 MW: 8C25AB66288ECAA4 CRC64;

Query Match 42.9%; Score 39; DB 1; Length 568;  
 Best Local Similarity 54.5%; Pred. No. 51;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 3 MYFSHRHAHLR 13  
 :||| ||| :  
 Db 210 LYFFHNHANFR 220  
 Search completed: November 27, 2002, 07:18:52  
 Job time : 13.1429 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 06:27:07 ; Search time 8.33977 Seconds  
(without alignments)  
370.599 Million cell updates/sec

Title: US-09-893-615-1

Perfect score: 91

Sequence: 1 WRMYFSHRRAHRLSP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_xvirus.\*

16: sp\_bacteriaph.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	51.6	757	2 Q9KIB0	Q9kib0 porphyromon
2	45	49.5	215	16 Q8ZCP9	Q8zcp9 yersinia pe
3	45	49.5	684	5 Q9VSL7	Q9vsl7 drosophila
4	45	49.5	684	10 Q9LZF4	Q9lzf4 arabisdopsis
5	44	48.4	307	16 Q34780	Q34780 bacillus su
6	44	48.4	703	10 Q32S84	Q32sb4 arabisdopsis
7	44	48.4	743	10 Q80707	Q80707 arabisdopsis
8	43.5	47.8	644	10 Q8S7K3	Q8s7k3 oryza sativ
9	43	47.3	421	10 Q9SB04	Q9sbu4 chlamydomon
10	43	47.3	424	10 Q48663	Q48663 chlamydomon
11	43	47.3	1358	10 Q23887	Q23887 zea mays (m
12	42	46.2	92	16 Q8XVZ0	Q8xvz0 ralstonia s
13	42	46.2	111	6 Q9N0J0	Q9n0j0 bos taurus
14	42	46.2	120	17 Q8TQU6	Q8tqu6 methanosarc
15	42	46.2	134	16 Q31649	Q31649 bacillus su
16	42	46.2	252	16 Q92DG9	Q92dg9 listeria in

17 42 46.2 252 16 Q8Y8P5 08y8p5 listeria mo  
18 42 46.2 486 5 Q9VQT3 Q9vqt3 drosophila  
19 42 46.2 879 4 Q9H3H4 Q9h3h4 homo sapien  
20 42 46.2 917 11 Q99K11 Q99k11 mus musculu  
21 42 46.2 935 5 Q17548 Q17548 caenorhabdi  
22 42 46.2 1435 12 Q9DSQ0 Q9dsq0 colorado ti  
23 42 46.2 1435 12 Q8QVF4 Q8qvf4 eyach virus  
24 42 46.2 2785 4 Q75691 Q75691 homo sapien  
25 41.5 45.6 248 2 Q9AJU8 Q9aju8 shigella fl  
26 41.5 45.6 267 2 Q8VSA9 Q8vsa9 shigella fl  
27 41.5 45.6 301 2 Q9AFK9 Q9afk9 shigella fl  
28 41 45.1 208 16 Q8YNI4 Q8yni4 anabaena sp  
29 41 45.1 228 16 Q9CNM1 Q9cnm1 pasteurella  
30 41 45.1 241 16 Q34901 Q34901 bacillus su  
31 41 45.1 285 16 Q9HV33 Q9hv33 pseudomonas  
32 41 45.1 309 5 Q9N9D3 Q9n9d3 physarum po  
33 41 45.1 310 5 Q9N9D2 Q9n9d2 physarum po  
34 41 45.1 346 2 Q44165 Q44165 actinobacil  
35 41 45.1 370 16 Q9A7P8 Q9a7p8 caulobacter  
36 41 45.1 378 10 Q942N6 Q942n6 oryza sativ  
37 41 45.1 452 4 Q96SP6 Q96sp6 homo sapien  
38 41 45.1 470 16 Q8YTC6 Q8ytc6 anabaena sp  
39 41 45.1 552 4 Q9NUG0 Q9nug0 homo sapien  
40 41 45.1 599 6 Q95LW5 Q95lw5 macaca fasc  
41 41 45.1 692 10 Q9LT62 Q9lt62 arabisdopsis  
42 41 45.1 822 10 Q9ZOX3 Q9zqx3 arabisdopsis  
43 41 45.1 833 12 Q69403 Q69403 pseudorabie  
44 41 45.1 838 12 Q65576 Q65576 bovine herp  
45 41 45.1 1063 10 Q80596 Q80596 arabisdopsis

## ALIGNMENTS

RESULT 1

Q9KIB0  
ID Q9KIB0 PRELIMINARY; PRT; 757 AA.  
AC Q9KIB0;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical tonB-linked outer membrane receptor PG13.  
OS Porphyromonas gingivalis (Bacteroides gingivalis).  
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;  
OC Porphyromonas.  
OX NCBI\_TaxID=837;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W50;  
RA Ross B., Barr J., Patterson M., Agius C., Rothel L., Margetts M.,  
RA Hocking D., Webb E.;  
RT "P. gingivalis polypeptides and nucleic acids."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W50;  
RA Ross B.C.;

RT "Genomic analysis of Porphyromonas gingivalis for vaccine discovery."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF237558; AAF81416.1; .  
DR InterPro; IPR00531; TonB\_boxC.  
DR Pfam; PF00593; TonB\_boxC; 1.  
KW Receptor.

SO SEQUENCE 757 AA; 84517 MW; 224B6D65264F9D62 CRC64;

Query Match 51.6%; Score 47; DB 2; Length 757;

Best Local Similarity 61.5%; Pred. NO.19;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WRMYFSHRRAHLR 13

Db 700 WRLATSHRFAHFR 712

```

RESULT 2
Q8ZCP9          PRELIMINARY;          PRT;      215 AA.
AC Q8ZCP9:
DT 01-NAR-2002 (TREMBLrel. 20, Created)
DT 01-NAR-2002 (TREMBLrel. 20, Last sequence update)
DE Putative membrane protein.
GN YPO2924.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT Nature 413:523-527(2001).
RL "Genome sequence of Yersinia pestis, the causative agent of plague."
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 215 AA; 24920 MW; EF312DBBAC15B3BB CRC64;

Query Match      49.5%; Score 45; DB 16; Length 215;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRNFSHRAHLR 13
   | | | | |
DB 72 WAMTGRRAHLR 84

RESULT 3
Q9VSL7          PRELIMINARY;          PRT;      684 AA.
AC Q9VSL7:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CG6817 protein.
GN FOI OR CG6817.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Boulc J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

```

```

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reihert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhou X., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003555; AAF50401.1; -.
DR FlyBase; FBgn024236; fol.
DR InterPro; IPR003689; Zn_trnprt.zip.
DR Pfam; PF02535; Zip. 1.
SQ SEQUENCE 684 AA; 75037 MW; A7272B67C342B988 CRC64;

Query Match      49.5%; Score 45; DB 5; Length 684;
Best Local Similarity 70.0%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 SHRHAHLRSP 15
   | | | | |
DB 505 SHRHGHVHSP 514

RESULT 4
Q9LZF4          PRELIMINARY;          PRT;      684 AA.
AC Q9LZF4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 79.7 kDa protein.
GN F12B4_90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162751; CAB83292.1; -.
DR InterPro; IPR004146; DCL.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03107; DCL. 3.
DR SMART; SM00184; RING. 2.
KW Hypothetical protein.
SQ SEQUENCE 684 AA; 79746 MW; C872720DD26E133E CRC64;

Query Match      49.5%; Score 45; DB 10; Length 684;
Best Local Similarity 70.0%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```



QY 4 YFSHRAHLR 13  
:||||:|  
DB 412 HFSHRHHLR 421

## RESULT 5

O34780 PRELIMINARY: PRT: 307 AA.  
AC O34780;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE RADC homolog.  
GN YOFB.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_taxid=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lapidus A., Galleron N., Sorokin A., Ehrlich D.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBDJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Gulseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosaoto V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the gram-positive bacterium Bacillus  
subtilis";  
RL Nature 390:249-256(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBDJ databases.  
DR EMBL; AF027868; AAB84459.1; -  
DR EMBL; Z99114; CAB13781.1; -  
KW Complete proteome.  
SQ SEQUENCE 307 AA; 35348 MW; BF3974AA51D0F762 CRC64;

Query Match 48.4%; Score 44; DB 16; Length 307;  
Best Local Similarity 58.3%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 WRMYFSHRAHLR 12  
:||||:|

DB 247 YAMYKDRHSHL 258

## RESULT 6

Q9ZSB4 PRELIMINARY: PRT: 703 AA.  
AC Q9ZSB4;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE F3H7.4 protein.  
GN F3H7.4 OR T4F9.20 OR AT4G10560.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA WASHU;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBDJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Laplant Y., Spalding L.;  
RL "The sequence of A. thaliana F3H7.";  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Waterston R.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBDJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Wedler H., Wedler E., Wambutt R., Bancroft I., Mewes H.W.,  
RA Mayer K.F.X., Schueller C.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBDJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBDJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBDJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RA Wedler H., Wedler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBDJ databases.  
RN [8]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBDJ databases.  
DR EMBL; AF118222; AAD03432.1; -  
DR EMBL; AL049523; CAB40022.1; -  
DR EMBL; AL161517; CAB78179.1; -  
DR InterPro; IPR004146; DCL.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF03107; DCL; 3.  
DR SMART; SM00184; RING; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 703 AA; 81396 MW; 6EC241534E870B41 CRC64;

Query Match 48.4%; Score 44; DB 10; Length 703;  
Best Local Similarity 70.0%; Pred. No. 54;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 4 YFSHRAHLR 13  
:||||:|  
DB 412 HFSHRHHLR 421

```
RESULT 7
O80707 PRELIMINARY; PRT; 743 AA.
AC O80707;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F8K4.22.
GN F8K4.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidops.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Kwan A., Yu G., Oji O.,
RA Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B.,
RA Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,
RA Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,
RA Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F8K4 sequence.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Theologis A.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004392; AAC28520.1; -.
DR InterPro; IPR002219; DAG_PE-bind.
DR Pfam; PF03107; DC1; 3.
DR SMART; SM00109; C1; 2.
SQ SEQUENCE 743 AA; 85956 MW; 7390A56812CC6F01 CRC64;

Query Match 48.4%; Score 44; DB 10; Length 743;
Best Local Similarity 70.0%; Pred. No. 57;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 YFSRHHLR 13
DB 418 HFSSHHLR 427

RESULT 8
Q8S7K3 PRELIMINARY; PRT; 644 AA.
AC Q8S7K3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative protein with similarity to protein phosphatase PPI regulatory
DE subunit.
GN OSJNB0095C06.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA Du H., Minx P., de la Bastide M., Nascimento L., Spiegel L.,
RA Preston R., Kirchoff K., Kuit K., Baker J., Vil M.D., Zutavern T.,
RA Santos L., Cunnius D.M., Rodriguez S., Miller B., Bailja V., Shah R.,
RA Bahret A., King L., O'Shaughnessy A., Palmer L., Dedhia N.,
RA McCombie W.R.;
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNB0095C06, from chromosome 10, complete sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079935; AAM08533.1; -.

Query Match 47.3%; Score 43; DB 10; Length 424;
Best Local Similarity 54.5%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAH 11
DB 176 WRIKHNNHHAH 186

RESULT 9
Q9SBU4 PRELIMINARY; PRT; 421 AA.
AC Q9SBU4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chloroplast w6 desaturase.
GN DES6.
OS Chlamydomonas sp. (strain W80).
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=103365;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W80;
RA Miyasaka H.;
RT "Cloning of chloroplast w6 desaturase of a marine halotolerant
RT Chlamydomonas";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB031546; BAA83822.1; -.
DR InterPro; IPR001225; FA_desaturase.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desaturase; 2.
SQ SEQUENCE 421 AA; 47932 MW; AD48F6039C8EDD57 CRC64;

Query Match 47.3%; Score 43; DB 10; Length 421;
Best Local Similarity 54.5%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAH 11
DB 176 WRIKHNNHHAH 186

RESULT 10
O48663 PRELIMINARY; PRT; 424 AA.
AC O48663;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chloroplast w6 desaturase.
GN DES6.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98158334; PubMed=9498569;
RX Sato N., Fujiwara S., Kawaguchi A., Tsuzuki M.;
RT "Cloning of a gene for chloroplast w6 desaturase of a green alga,
RT Chlamydomonas reinhardtii.";
RL J. Biochem. 122:1224-1232(1997).
DR EMBL; AB007640; BAA23881.1; -.
DR InterPro; IPR001225; FA_desaturase.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desaturase; 2.
SQ SEQUENCE 424 AA; 48377 MW; D553054C1CEE6B0D CRC64;

Query Match 47.3%; Score 43; DB 10; Length 424;
```

```
Best Local Similarity 54.5%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 WRMYFSRHRAH 11
DB 178 WRKHNHHHAH 188
      ||: :|||
      ||: :|||

RESULT 11
O23887 PRELIMINARY; PRT; 1358 AA.
AC O23887
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Aldehyde oxidase (EC 1.2.3.1).
GN ZMAO-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLEOPTILE;
RX MEDLINE=97326103; PubMed=9182554;
RA Sekimoto H., Seo M., Dohmae N., Takio K., Kamiya Y., Koshiba T.;
RT "Cloning and molecular characterization of plant aldehyde oxidase.";
RL J. Biol. Chem. 272:15280-15285(1997).
CC -1- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
DR EMBL; D88451; BAA23236.1; -.
DR HSSP; P80457; 1FTQ.
DR InterPro; IPR002888; 2Fe-2S_bind.
DR InterPro; IPR000564; 2Fe2S_ferredoxin.
DR InterPro; IPR000674; Aldxan_dh_C.
DR InterPro; IPR005107; CO_doh_flav_C.
DR InterPro; IPR002346; dehydrog_molyb.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001041; Ferredoxin.
DR Pfam; PF01315; Ald_xan_dh_C; 1.
DR Pfam; PF02738; Ald_xan_dh_C2; 1.
DR Pfam; PF03450; CO_doh_flav_C; 1.
DR Pfam; PF00941; FAD_binding_5; 1.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF01799; fer2.2; 1.
DR ProDom; PD186071; 2Fe-2S_bind; 1.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW Iron-sulfur; Oxidoreductase.
SQ SEQUENCE 1358 AA; 146682 MW; 82449227AFB14861 CRC64;

Query Match 47.3%; Score 43; DB 10; Length 1358;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 MYFSRHRAHLS 14
DB 619 IYSTHPHVRFS 630
      ||: |||:|
      ||: |||:|

RESULT 12
O8XV20 PRELIMINARY; PRT; 92 AA.
AC O8XV20
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein RSC2685.
GN RSC2685 OR RS00036.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
```

```
RN SEQUENCE FROM N.A.
RP STRAIN-CM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646071; CAD16392.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 92 AA; 10226 MW; DB684B409C8FC62B CRC64;

Query Match 46.2%; Score 42; DB 16; Length 92;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 WRMYFSRHRAHLS 14
DB 30 WLQFFGDRHGFLRS 43
      | : | | | | |
      | : | | | | |

RESULT 13
Q9NOJ0 PRELIMINARY; PRT; 111 AA.
AC Q9NOJ0
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Angiotensin receptor 2 (Fragment).
GN AT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Berisha B.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RX MEDLINE=20039923; PubMed=10569984;
RA Acosta T., Berisha B., Ozawa T., Sato K., Schams D., Miyamoto A.;
RT "Evidence for local endothelin-angiotensin atrial natriuretic peptide
RT system in bovine mature follicles in vitro: effects on steroid
RT hormones and prostaglandin secretion.";
RL Biol. Reprod. 61:1419-1425(1999).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AJ277986; CAB92128.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
FT NON_TER 1
FT NON_TER 111
SQ SEQUENCE 111 AA; 13102 MW; 9E8A81D3D580A601 CRC64;

Query Match 46.2%; Score 42; DB 6; Length 111;
Best Local Similarity 46.7%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WRMYFSRHRAHLS 15
DB 7 WATYSHRYDWLFGP 21
      | : | | | | |
      | : | | | | |
```

## RESULT 14

Q8TQ06 PRELIMINARY; PRT; 120 AA.  
 AC Q8TQ06; 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Predicted protein.  
 GN M1443.  
 OS Methanosarcina acetivorans.  
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;  
 OC Methanosarcinaceae; Methanosarcina.  
 OX NCBI\_TaxID=2214;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
 RX MEDLINE=21929760; PubMed=11932238;  
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,  
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,  
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
 RA Hederich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,  
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
 RA Metcalf W.W., Birren B.;  
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
 RT and physiological diversity."  
 RL Genome Res. 12:532-542(2002).  
 DR EMBL: AE010814; AAM04857.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 120 AA; 13232 MW; 49AAB649268558FD CRC64;

Query Match 46.2%; Score 42; DB 17; Length 120;  
 Best Local Similarity 72.7%; Pred. No. 20;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 FSHRHAHLRSP 15  
 ||||| | | |

Db 75 FSHRYAQLSP 85

## RESULT 15

Q31649 PRELIMINARY; PRT; 134 AA.  
 AC Q31649;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE YJH protein.  
 GN YJH.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Key M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viarl A., Wambutt R., Wedler E., Wedler H., Weitzsaecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the gram-positive bacterium Bacillus  
 RT subtilis."  
 RL Nature 390:249-256(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z99110; CAB13062.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 134 AA; 15340 MW; 80F1E6CFC76B34E3 CRC64;

Query Match 46.2%; Score 42; DB 16; Length 134;  
 Best Local Similarity 66.7%; Pred. No. 22;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 YFSHRHAHLRSP 15  
 ||||| | | |

Db 21 YFSHTNPQRSP 32

Search completed: November 27, 2002, 07:22:39

Job time : 21.3398 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 07:17:51 ; Search time 2.72201 Seconds  
(without alignments)  
162.139 Million cell updates/sec

Title: US-09-893-615-1

Perfect score: 91

Sequence: 1 WRMYFSHRHAHLRSP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	44.0	264	4	US-09-069-023-7
2	40	44.0	300	4	US-09-099-041A-4
3	40	44.0	300	4	US-09-245-281-4
4	40	44.0	300	4	US-09-207-359B-4
5	40	44.0	531	4	US-09-069-023-1
6	40	44.0	540	3	US-09-018-942-1
7	40	44.0	540	4	US-09-099-041A-2
8	40	44.0	540	4	US-09-069-023-27
9	40	44.0	540	4	US-09-245-281-2
10	40	44.0	540	4	US-09-470-271-1
11	40	44.0	540	4	US-09-207-359B-2
12	39	42.9	1019	1	US-08-271-364A-7
13	39	42.9	1019	2	US-08-222-715B-26
14	38	41.8	305	4	US-09-134-001C-5632
15	37	40.7	29	4	US-09-354-231B-34
16	37	40.7	29	4	US-09-128-602B-34
17	37	40.7	189	2	US-08-837-029-2
18	37	40.7	193	2	US-08-336-031-6
19	37	40.7	193	5	PCT-US95-06725-6
20	37	40.7	224	4	US-08-530-862B-11
21	37	40.7	224	4	US-08-597-313D-11
22	37	40.7	224	4	US-09-133-962A-10
23	37	40.7	355	3	US-08-872-979-7
24	37	40.7	396	1	US-08-229-284A-2
25	37	40.7	411	2	US-08-336-031-2
26	37	40.7	411	2	US-08-902-853-7
27	37	40.7	411	5	PCT-US95-06725-2

28	37	40.7	774	3	US-08-902-632-2	Sequence 2, Appl
29	37	40.7	774	3	US-09-073-354-1	Sequence 1, Appl
30	37	40.7	774	3	US-08-656-005A-1	Sequence 1, Appl
31	37	40.7	774	4	US-09-073-259-1	Sequence 1, Appl
32	37	40.7	774	4	US-09-363-095-1	Sequence 1, Appl
33	37	40.7	774	4	US-09-418-027-1	Sequence 1, Appl
34	37	40.7	921	4	US-09-206-800-9	Sequence 9, Appl
35	37	40.7	1164	2	US-08-589-756-1	Sequence 1, Appl
36	37	40.7	1164	4	US-09-206-800-1	Sequence 1, Appl
37	37	40.7	1164	4	US-09-206-898-1	Sequence 1, Appl
38	37	40.7	1313	4	US-08-989-299-9	Sequence 9, Appl
39	36.5	40.1	840	4	US-08-974-549A-190	Sequence 190, App
40	36.5	40.1	872	3	US-08-851-843A-8	Sequence 8, Appl
41	36.5	40.1	872	3	US-08-851-843A-54	Sequence 54, Appl
42	36.5	40.1	872	4	US-08-974-549A-221	Sequence 221, App
43	36.5	40.1	872	4	US-08-854-050-8	Sequence 8, Appl
44	36.5	40.1	872	4	US-08-854-050-54	Sequence 54, Appl
45	36.5	40.1	872	4	US-09-430-323-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1  
US-09-069-023-7  
; Sequence 7, Application US/09069023A  
; Patent No. 6348573  
; GENERAL INFORMATION:  
; APPLICANT: Nunez, Gabriel  
; APPLICANT: Inohara, Naohiro  
; APPLICANT: Koseki, Takeyoshi  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS  
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS  
; FILE REFERENCE: UM-03333  
; CURRENT APPLICATION NUMBER: US/09/069,023A  
; CURRENT FILING DATE: 1998-04-27  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 264  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-069-023-7

Query Match 44.0%; Score 40; DB 4; Length 264;  
Best Local Similarity 33.3%; Pred. No. 28;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WRMYFSHRHAHLRSP 15  
||: :| |:|  
Db 30 WRVQVAVKHLHTWP 44

RESULT 2  
US-09-099-041A-4  
; Sequence 4, Application US/09099041A  
; Patent No. 6340576  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 07334-076001  
; CURRENT APPLICATION NUMBER: US/09/099,041A  
; CURRENT FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 09/019,942  
; PRIOR FILING DATE: 1998-02-06  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-099-041A-4

Query Match 44.0%; Score 40; DB 4; Length 300;  
Best Local Similarity 33.3%; Pred. No. 32;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
||: : || : ||  
DB 40 WRVQVAVKHLHIHP 54

## RESULT 3

US-09-245-281-4  
; Sequence 4, Application US/09245281

; Patent No. 6369196

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY

; FILE REFERENCE: 07334/118001

; CURRENT APPLICATION NUMBER: US/09/245,281

; CURRENT FILING DATE: 1999-02-05

; EARLIER APPLICATION NUMBER: US 09/207,359

; EARLIER FILING DATE: 1998-12-08

; EARLIER APPLICATION NUMBER: US 09/099,041

; EARLIER FILING DATE: 1998-06-17

; EARLIER APPLICATION NUMBER: US 09/019,942

; EARLIER FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 300

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-245-281-4

Query Match 44.0%; Score 40; DB 4; Length 300;  
Best Local Similarity 33.3%; Pred. No. 32;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
||: : || : ||  
DB 40 WRVQVAVKHLHIHP 54

## RESULT 4

US-09-207-359B-4

; Sequence 4, Application US/09207359B

; Patent No. 6469140

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE REFERENCE: 07334-112001

; CURRENT APPLICATION NUMBER: US/09/207,359B

; CURRENT FILING DATE: 1998-12-08

; PRIOR APPLICATION NUMBER: US 09/099,041

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: US 09/019,942

; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 300

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-207-359B-4

Query Match 44.0%; Score 40; DB 4; Length 300;  
Best Local Similarity 33.3%; Pred. No. 32;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
||: : || : ||

Db 40 WRVQVAVKHLHIHP 54

## RESULT 5

US-09-069-023-1

; Sequence 1, Application US/09069023A

; Patent No. 6348573

; GENERAL INFORMATION:

; APPLICANT: Nunez, Gabriel

; APPLICANT: Inohara, Naohiro

; APPLICANT: Koseki, Takeyoshi

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

; FILE REFERENCE: UM-03333

; CURRENT APPLICATION NUMBER: US/09/069,023A

; CURRENT FILING DATE: 1998-04-27

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 531

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-069-023-1

Query Match 44.0%; Score 40; DB 4; Length 531;  
Best Local Similarity 33.3%; Pred. No. 58;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
||: : || : ||  
DB 31 WRVQVAVKHLHIHP 45

## RESULT 6

US-09-019-942-1

; Sequence 1, Application US/09019942

; Patent No. 6033855

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT

; FILE REFERENCE: 07334-112001

; CURRENT APPLICATION NUMBER: US/09/019,942

; CURRENT FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: US 09/019,942

; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 2.0b

; SEQ ID NO 4

; LENGTH: 300

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-019-942-1

Query Match 44.0%; Score 40; DB 4; Length 300;  
Best Local Similarity 33.3%; Pred. No. 58;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRHAHLRSP 15  
||: : || : ||  
DB 31 WRVQVAVKHLHIHP 45

## RESULT 6

US-09-019-942-1

; Sequence 1, Application US/09019942

; Patent No. 6033855

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT

; FILE REFERENCE: 07334-112001

; CURRENT APPLICATION NUMBER: US/09/019,942

; CURRENT FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: US 09/019,942

; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 2.0b

; SEQ ID NO 4

; LENGTH: 300

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-019-942-1

; MOLECULE TYPE: protein  
US-09-019-942-1

Query Match 44.0%; Score 40; DB 3; Length 540;  
Best Local Similarity 33.3%; Pred. No. 59;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WRMYFSHRHAHLRSP 15  
||: : | | : |  
Db 40 WRVQVAVKHLHIHTP 54

## RESULT 7

US-09-099-041A-2  
; Sequence 2, Application US/09099041A  
; Patent No. 6340576

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE REFERENCE: 07334-076001

; CURRENT APPLICATION NUMBER: US/09/099,041A

; CURRENT FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 09/019,942

; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 540

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-099-041A-2

Query Match 44.0%; Score 40; DB 4; Length 540;  
Best Local Similarity 33.3%; Pred. No. 59;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WRMYFSHRHAHLRSP 15  
||: : | | : |  
Db 40 WRVQVAVKHLHIHTP 54

## RESULT 8

US-09-069-023-27

; Sequence 27, Application US/09069023A

; Patent No. 6348573

; GENERAL INFORMATION:

; APPLICANT: Nunez, Gabriel

; APPLICANT: Inohara, Naohiro

; APPLICANT: Koseki, Takeyoshi

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

; FILE REFERENCE: UM-03333

; CURRENT APPLICATION NUMBER: US/09/069,023A

; CURRENT FILING DATE: 1998-04-27

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 27

; LENGTH: 540

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-069-023-27

Query Match 44.0%; Score 40; DB 4; Length 540;  
Best Local Similarity 33.3%; Pred. No. 59;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WRMYFSHRHAHLRSP 15  
||: : | | : |  
Db 40 WRVQVAVKHLHIHTP 54

## RESULT 9

US-09-245-281-2  
; Sequence 2, Application US/09245281  
; Patent No. 6369196  
; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY

; FILE REFERENCE: 07334/118001

; CURRENT APPLICATION NUMBER: US/09/245,281

; EARLIER FILING DATE: 1998-02-05

; EARLIER FILING DATE: 1998-12-08

; EARLIER FILING DATE: 1998-06-17

; EARLIER FILING DATE: 1998-06-17

; EARLIER FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 540

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-245-281-2

Query Match 44.0%; Score 40; DB 4; Length 540;  
Best Local Similarity 33.3%; Pred. No. 59;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WRMYFSHRHAHLRSP 15  
||: : | | : |  
Db 40 WRVQVAVKHLHIHTP 54

## RESULT 10

US-09-470-271-1

; Sequence 1, Application US/09470271

; Patent No. 6410689

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: FastSeq for Windows Version 2.0b

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/470,271

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/019,942

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Melkjohn, Ph.D., Anita L.

; REGISTRATION NUMBER: 35,283

; REFERENCE/DOCKET NUMBER: 07334/068001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 540 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein  
US-09-470-271-1

Query Match 44.0%; Score 40; DB 4; Length 540;  
Best Local Similarity 33.3%; Pred. No. 59;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRHRLRSP 15  
|| : : || : :  
DB 40 WRVQVAVKHLHIHTP 54

## RESULT 11

US-09-207-359B-2

; Sequence 2, Application US/09207359B

; Patent No. 6469140

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE REFERENCE: 07334-112001

; CURRENT APPLICATION NUMBER: US/09/207,359B

; CURRENT FILING DATE: 1998-12-08

; PRIOR APPLICATION NUMBER: US 09/099,041

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: US 09/019,942

; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 540

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-207-359B-2

Query Match 44.0%; Score 40; DB 4; Length 540;  
Best Local Similarity 33.3%; Pred. No. 59;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMYFSHRHRLRSP 15  
|| : : || : :  
DB 40 WRVQVAVKHLHIHTP 54

## RESULT 12

US-08-271-364A-7

; Sequence 7, Application US/08271364A

; Patent No. 5756334

; GENERAL INFORMATION:

; APPLICANT: PERLER, FRANCINE B.

; APPLICANT: SOUTHWORTH, MAURICE W.

; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE DNA POLYMERASE

; FROM ARCHAEABACTERIA

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NEW ENGLAND BIOLABS, INC.

; STREET: 32 TOZER ROAD

; CITY: BEVERLY

; STATE: MASSACHUSETTS

; COUNTRY: US

; ZIP: 01915

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/271,364A

; FILING DATE: 06-JUL-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/811,421

; FILING DATE: 18-DEC-1991

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/686,340  
; FILING DATE: 17-APR-1991

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/626,057

; FILING DATE: 11-DEC-1990

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/513,994

; FILING DATE: 26-APR-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: WILLIAMS, GREGORY D.

; REGISTRATION NUMBER: 30901

; REFERENCE/DOCKET NUMBER: NEB-101

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508) 927-5054

; TELEFAX: (508) 927-1705

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1019 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-08-271-364A-7

Query Match 42.9%; Score 39; DB 1; Length 1019;  
Best Local Similarity 71.4%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WRMYFSH 7  
|| : : || : :  
DB 83 WRLYPEH 89

## RESULT 13

US-08-222-715B-26

; Sequence 26, Application US/08222715B

; Patent No. 5834285

; GENERAL INFORMATION:

; APPLICANT: Comb, Donald G.

; APPLICANT: Perler, Francine

; APPLICANT: Kucera, Rebecca

; APPLICANT: Jack, William E.

; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE DNA

; POLYMERASE FROM ARCHAEABACTERIA

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS,

; INC.

; STREET: 32 TOZER ROAD

; CITY: BEVERLY

; STATE: MASSACHUSETTS

; COUNTRY: US

; ZIP: 01915

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/222,715B

; FILING DATE: 04-APR-1994

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/167,238

; FILING DATE: 15-DEC-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/686,340

; FILING DATE: 17-APR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/626,057

; FILING DATE: 11-DEC-1990

; PRIOR APPLICATION DATA:



; APPLICATION NUMBER: US 07/513,994  
; FILING DATE: 26-APR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams, Gregory D.  
; REGISTRATION NUMBER: 30901  
; REFERENCE/DOCKET NUMBER: NEB-054C3FC2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508) 927-5054  
; TELEFAX: (508) 927-1705  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1019 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-222-715B-26

Query Match 42.9%; Score 39; DB 2; Length 1019;  
Best Local Similarity 71.4%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WRMYFSH 7  
|||:|  
Db 83 WRLYFEH 89

## RESULT 14

US-09-134-001C-5632  
; Sequence 5632, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134.001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5632  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5632

Query Match 41.8%; Score 38; DB 4; Length 305;  
Best Local Similarity 42.9%; Pred. No. 70;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WRMYFSHRHAHLRS 14  
: :| |||:  
Db 289 YNVVYVHEDAHIES 302

## RESULT 15

US-09-354-231B-34  
; Sequence 34, Application US/09354231B  
; Patent No. 6342658  
; GENERAL INFORMATION:  
; APPLICANT: DeBonte, Lorin R.  
; APPLICANT: Shorosh, Basil S.  
; TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF  
; FILE REFERENCE: 07148-063002  
; CURRENT APPLICATION NUMBER: US/09/354,231B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: US 08/874,109  
; PRIOR FILING DATE: 1997-06-12  
; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Ricinus communis  
US-09-354-231B-34

Query Match 40.7%; Score 37; DB 4; Length 29;  
Best Local Similarity 60.0%; Pred. No. 9.2;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 YFSHRHAHLR 13  
|||:|  
Db 5 YFSWKHSHRR 14

Search completed: November 27, 2002, 07:25:10  
Job time : 5.72201 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:38:32 ; Search time 67.9112 Seconds  
(without alignments)  
241.342 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 646

Sequence: 1 EYLVESGGGLVQPKGSLK.....SGIDYAMYWGQGTSLTVSS 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555	85.9	115	AAW01593	Lead binding MAB 2
2	549	85.0	119	AAW08029	3B1 heavy chain va
3	549	85.0	599	AAW08037	3B1 single chain a
4	536	83.0	333	AA170111	Anti-Her2neu singl
5	532	82.4	110	AAW01587	Lead binding MAB 6
6	519	80.3	117	AAW09919	Murine mab lD9 hea
7	519	80.3	117	AAW06947	Murine lD9 antibod
8	519	80.3	148	AAW07033	Murine antibody lD
9	514.5	79.6	125	AA110022	H. pylori 26 kDa p
10	510	78.9	117	AAW014978	Mouse heavy chain

11	493	76.3	117	22	AAU09930	Humanised lD9 heav
12	493	76.3	117	22	AAE06957	Humanised murine l
13	489	75.7	117	22	AAU09928	Humanised lD9 heav
14	489	75.7	117	22	AAU09929	Humanised lD9 heav
15	489	75.7	117	22	AAE06956	Humanised murine l
16	489	75.7	117	23	AAQ14982	Humanised murine h
17	488	75.5	117	22	AAE06955	Humanised murine h
18	488	75.5	117	23	AAQ14981	Humanised murine h
19	487	75.4	117	22	AAU09927	Humanised lD9 heav
20	487	75.4	117	22	AAE06954	Humanised murine l
21	487	75.4	117	23	AAQ14980	Humanised murine h
22	487	75.4	119	22	AAE07034	Humanised murine a
23	483	74.8	101	22	AAE06971	Murine lD9 antibod
24	482.5	74.7	118	23	AAQ14983	Humanised murine h
25	476	73.7	119	19	AAW46958	Amino acid sequenc
26	471.5	73.0	126	22	AAE07011	Human heavy chain
27	468	72.4	100	22	AAE06972	Mouse germline hea
28	467.5	72.4	256	23	AAU72866	P5-3 single chain
29	467.5	72.4	503	23	AAU72874	3B10P5-23 bispec1
30	463	71.7	127	22	AAE07008	Human heavy chain
31	462.5	71.6	122	21	AAW90812	2G3 hybridoma VH d
32	460.5	71.3	120	18	AAW27553	Human Ab heavy cha
33	460.5	71.3	281	18	AAW27560	Consensus single c
34	456	70.6	123	22	AAU02615	Anti-adipocyte mon
35	455.5	70.5	126	22	AAE07010	Human heavy chain
36	454.5	70.4	120	18	AAW01589	Lead binding MAB 4
37	451.5	69.9	122	21	AAV50973	Human FVIII antibo
38	451.5	69.9	122	21	AAV50975	Human FVIII heavy
39	448.5	69.4	444	21	AAV32263	Humanised anti-CD2
40	448	69.3	121	20	AAW86122	Protein sequence o
41	448	69.3	570	20	AAV39451	Antibody ABX-CBL h
42	447.5	69.3	120	19	AAW23952	Chimeric humanised
43	447.5	69.3	240	19	AAW23954	Chimeric humanised
44	447	69.2	116	20	AAV03869	SM3 heavy chain va
45	447	69.2	142	18	AAW06212	MAB Br-3 heavy cha

ALIGNMENTS

- RESULT 1
- AAW01593
- ID AAW01593 standard; Protein; 115 AA.
- XX
- AC AAW01593;
- XX
- DT 22-AUG-1997 (first entry)
- XX
- DE Lead binding MAB 2B4 heavy chain variable region.
- XX
- KW Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic;
- KW pharmaceutical; health care; skin treatment; pesticide; herbicide;
- KW heavy metal.
- XX
- OS Mus musculus.
- XX
- PN WO9639518-A1.
- XX
- PD 12-DEC-1996.
- XX
- PF 05-JUN-1996; 96WO-US09258.
- XX
- PR 10-OCT-1995; 95US-0541373.
- XX
- PR 05-JUN-1995; 95US-0462798.
- XX
- PA (BION-) BIONEBRASKA INC.
- XX
- PI Lopez O, Murray PJ, Wylie DE;
- XX
- XX WPI; 1997-043140/04.
- DR N-PSDB; AAT58267.
- XX
- XX DNA encoding heavy metal binding polypeptide sequences - used for
- PT

PT detecting, removing, adding or neutralising heavy metals, such as  
 XX lead cations  
 PS Claim 12; Page 89; 125pp; English.  
 XX The present sequence represents the heavy chain variable region for  
 CC monoclonal antibody (MAb) 2B4, which immunoreacts with a lead cation.  
 CC The sequence was derived from RNA isolated from mouse hybridoma cells.  
 CC The protein can be used for binding heavy metals, such as lead cations.  
 CC It can be used for detecting, removing, adding or neutralising the  
 CC heavy metals in biological and inanimate systems. It can be used in  
 CC e.g. aqueous liquid systems, in biological or environmental systems or  
 CC in such compositions as perfumes, cosmetics, pharmaceuticals, health  
 CC care products, skin treatment products, pesticides, herbicides,  
 CC solvents used in the production of semi-conductor and integrated  
 CC circuit components and production materials for electronic components.  
 CC The products can provide for applications involving minute amounts of  
 CC specific heavy metals.

XX Sequence 115 AA;

Query Match 85.9%; Score 555; DB 18; Length 115;

Best Local Similarity 87.8%; Pred. No. 2.6e-46;

Matches 108; Conservative 2; Mismatches 5; Indels 8; Gaps 1;

QY 1 EYLVESGGGLVQPKGSLKLSAASGFTFNMYAMNVRQAPGKGLEWVARIRSKSNYYAT 60

DB 1 EVQLVESGGGLVQPKGSLKLSAASGFTFNMYAMNVRQAPGKGLEWVARIRSKSNYYAT 60

QY 61 FYADSVKDRFTISRDDSQSMFLQMNKLKTEDTAMYYCVRGASGIDYAMDYWGQGTSLT 120

DB 61 YYADSVKDRFTISRDDSQSMFLQMNKLKTEDTAMYYCVR-----RDYWGQGTSLT 112

QY 121 VSS 123

DB 113 VSS 115

RESULT 2

AAR90829

ID AAR90829 standard; Protein; 119 AA.

XX AAR90829;

XX 25-JUN-1996 (first entry)

DE 3B1 heavy chain variable region from pCIB4613.

XX delta endotoxin; Bacillus thuringiensis; western corn rootworm;

KW WCRW; maize; pesticide; brush border membrane vesicle; monoclonal;

KW antibody.

XX Insecta sp.

XX WO9600783-A1.

PN 11-JAN-1996.

PD 20-JUN-1995; 95WO-IB00497.

PF 28-JUN-1994; 94US-0267641.

XX (CIBA ) CIBA GEIGY AG.

XX Carozzi NB, Koziel MG;

XX WPI; 1996-077494/08.

DR N-PSDB; AAT15725.

XX New monoclonal antibodies which bind insect gut proteins - used  
 PT partic. with toxin moieties for the control of insect pests, partic.  
 PT in plants

XX

PS Claim 8; Page 50-51; 106pp; English.

XX AAR90829-39 are monoclonal antibodies or a binding fragments produced by  
 CC using insect guts, partic. insect brush border membranes (BBMs), esp.  
 CC corn rootworm, as antigen; immunising a donor animal with the antigen;  
 CC isolating immunocompetent B cells from the immunised animal; fusing B  
 CC cells with a tumour cell line; isolating the fused cells, culturing them  
 CC and cloning positive hybrid cells; and screening the hybrid cells for  
 CC prodn. of the required MABs. The MABs bind to the gut of a target insect  
 CC but do not bind to mammalian BBMs. The DNA sequence can be operably  
 CC linked to a toxin moiety, esp. selected from e.g. Bacillus toxins,  
 CC Pseudomonas exotoxin and phytolectin, etc.. The Abs are useful for  
 CC control of insect pests, e.g. Coleoptera, Diptera, Hymenoptera and  
 CC Lepidoptera. The pesticidal compsn. is pref. applied to a plant, e.g.  
 CC maize.

XX Sequence 119 AA;

Query Match 85.0%; Score 549; DB 17; Length 119;

Best Local Similarity 86.3%; Pred. No. 1e-45;

Matches 107; Conservative 5; Mismatches 6; Indels 6; Gaps 2;

QY 1 EYLVESGGGLVQPKGSLKLSAASGFTFNMYAMNVRQAPGKGLEWVARIRSKSNYYAT 60

DB 1 QVQLQESGGGLVQPKGSLKLSAASGFTFNFMNVRQAPGKGLEWVARIRSKSNYYAT 60

QY 61 FYADSVKDRFTISRDDSQSMFLQMNKLKTEDTAMYYCVRGASGIDY-AMDYWGQGTSL 119

DB 61 SYGDSVKDRFTISRDDSQSMFLQMNKLKTEDTAMYYCVR-----VVYGDYWGQGTSL 115

QY 120 TVSS 123

DB 116 TVSS 119

RESULT 3

AAR90837

ID AAR90837 standard; Protein; 599 AA.

XX AAR90837;

XX 25-JUN-1996 (first entry)

DE 3B1 single chain antibody from pCIB4631.

XX delta endotoxin; Bacillus thuringiensis; western corn rootworm;

KW WCRW; maize; pesticide; brush border membrane vesicle; monoclonal;

KW antibody.

XX Insecta sp.

XX WO9600783-A1.

PN 11-JAN-1996.

PF 20-JUN-1995; 95WO-IB00497.

PR 28-JUN-1994; 94US-0267641.

XX (CIBA ) CIBA GEIGY AG.

XX Carozzi NB, Koziel MG;

XX WPI; 1996-077494/08.

DR N-PSDB; AAT15733.

XX New monoclonal antibodies which bind insect gut proteins - used  
 PT partic. with toxin moieties for the control of insect pests, partic.  
 PT in plants

PS Claim 8; Page 68-72; 106pp; English.

XX AAR90829-39 are monoclonal antibodies or a binding fragments produced by

XX

CC using insect guts, partic. insect brush border membranes (BBMs), esp.  
CC corn rootworm, as antigen; immunising a donor animal with the antigen;  
CC isolating immunocompetent B cells from the immunised animal; fusing B  
CC cells with a tumour cell line; isolating the fused cells, culturing them  
CC and cloning positive hybrid cells; and screening the hybrid cells for  
CC prodn. of the required MABs. The MABs bind to the gut of a target insect  
CC but do not bind to mammalian BBMs. The DNA sequence can be operably  
CC linked to a toxin moiety, esp. selected from e.g. Bacillus toxins,  
CC Pseudomonas exotoxin and phytoalexin, etc. The Abs are useful for  
CC control of insect pests, e.g. Coleoptera, Diptera, Hymenoptera and  
CC Lepidoptera. The pesticidal compsn. is pref. applied to a plant, e.g.  
CC maize.

XX SQ Sequence 599 AA;

Query Match 85.0%; Score 549; DB 17; Length 599;  
Best Local Similarity 86.3%; Pred. No. 6.6e-45;  
Matches 107; Conservative 5; Mismatches 6; Indels 6; Gaps 2;

Qy 1 EYMLVESGGGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPGKGLWVARIRKSNYYAT 60  
Db 155 QVWLESGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPGKGLWVARIRKSNYYAT 214  
Qy 61 FYADSVKDRFTISRDDQSMLYLQNNLKTEDTAMYYCVRGASGIDY-AMDYWGQGTSL 119  
Db 215 SYGDSVKDRFTISRDDQSMSFYLQNNLKTEDTAMYYCVR----VYIGAMDYWGQGTSLV 269  
Qy 120 TVSS 123  
Db 270 TVSS 273

#### RESULT 4

AAW01587  
ID AAW01587 standard; Protein: 333 AA.

XX AC AAW01587;

XX DT 05-JUN-2000 (first entry)

XX DE Anti-Her2neu single chain antibody.

XX KW Retroviral vector; cell specific gene transfer; Spleen Necrosis Virus;  
XX SNV; antibody-envelope fusion protein; retroviral envelope protein;  
XX gene therapy; antigen binding site; single chain antibody; human;  
XX Her2neu cell surface marker.

XX OS Homo sapiens.

XX PN WO200009730-A2.

XX PD 24-FEB-2000.

XX PF 10-AUG-1999; 99WO-US18141.

XX PR 17-AUG-1998; 98US-0135121.

XX PA (UYJE-) UNIV JEFFERSON THOMAS.

XX PI Dornburg RC;

XX DR WPI; 2000-224358/19.

XX DR N-PSDB; AAZ51115.

XX PT Cell specific gene transfer using retroviral vectors containing  
XX antibody-envelope fusion proteins and wild type envelope proteins -

XX PS Example; Fig 8; 45pp; English.

XX CC The patent discloses a novel retroviral vector, particularly Spleen  
XX Necrosis Virus (SNV) vector, having target cell specificity. The vector  
XX has a targetting envelope which is a chimeric protein consisting of an  
XX antigen binding site of an antibody (e.g. anti-DNP-scfv) or another

CC peptide that binds to a specific cell surface protein, fused to the  
CC carboxy terminal part of the retroviral envelope protein. The presence  
CC of the wild type envelope protein serves as a helper molecule to  
CC improve or supplement a functional membrane fusion domain. The  
CC antigen binding site replaces the natural viral receptor binding site.  
CC The retroviral vector is used for cell specific gene transfer,  
CC especially in gene therapy. The invention overcomes the restricted host  
CC range limitation of retroviral vectors. The present sequence is an  
CC anti-Her2neu single chain antibody (sca). The coding sequence of  
CC this antibody was used in the construction of plasmid pAJ7 which  
CC contains a targetting vector comprising the anti-Her2neu sca fused to the  
CC SNV-Env(envelope)-TM(transmembrane) coding region.

XX SQ Sequence 333 AA;

Query Match 83.0%; Score 536; DB 21; Length 333;  
Best Local Similarity 83.3%; Pred. No. 6.2e-44;  
Matches 105; Conservative 4; Mismatches 11; Indels 6; Gaps 2;

Qy 1 EYMLVESGGGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPGKGLWVARIRKSNYYAT 60  
Db 144 EYMLVESGGGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPGKGLWVARIRKSNYYAT 203  
Qy 61 FYADSVKDRFTISRDDQSMLYLQNNLKTEDTAMYYCVRGASGIDY-AMDYWGQGT 117  
Db 204 YIVDSVKDRFTISRDDQSMLYLQNNLKTEDTAMYYCVR---TSYYDYDKVLFAWGQGT 260  
Qy 118 SLTVSS 123  
Db 261 TVTVSS 266

#### RESULT 5

AAW01587  
ID AAW01587 standard; Protein: 110 AA.

XX AC AAW01587;

XX DT 22-AUG-1997 (first entry)

XX DE Lead binding MAB 6F5 heavy chain variable region.

XX KW Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic;  
XX pharmaceutical; health care; skin treatment; pesticide; herbicide;  
XX heavy metal.

XX OS Mus musculus.

XX PN WO9639518-A1.

XX PD 12-DEC-1996.

XX PF 05-JUN-1996; 96WO-US09258.

XX PR 10-OCT-1995; 95US-0541373.

XX PR 05-JUN-1995; 95US-0462798.

XX PA (BION-) BIONEERASKA INC.

XX PI Lopez O, Murray PJ, Wylie DE;

XX DR WPI; 1997-043140/04.

XX DR N-PSDB; AAT58261.

XX PT DNA encoding heavy metal binding polypeptide sequences - used for  
XX detecting, removing, adding or neutralising heavy metals, such as  
XX lead cations

XX PS Claim 12; Page 77; 125pp; English.

XX CC The present sequence represents the heavy chain variable region for  
XX monoclonal antibody (MAB) 6F5, which immunoreacts with a lead cation.  
XX The sequence was derived from RNA isolated from mouse hybridoma cells.

CC The protein can be used for binding heavy metals, such as lead cations.  
CC It can be used for detecting, removing, adding or neutralising the  
CC heavy metals in biological and inanimate systems. It can be used in  
CC e.g. aqueous liquid systems, in biological or environmental systems or  
CC in such compositions as perfumes, cosmetics, pharmaceuticals, health  
CC care products, skin treatment products, pesticides, herbicides,  
CC solvents used in the production of semi-conductor and integrated  
CC circuit components and production materials for electronic components.  
CC The products can provide for applications involving minute amounts of  
CC specific heavy metals.

XX SQ Sequence 110 AA;  
Query Match 82.4%; Score 532; DB 18; Length 110;  
Best Local Similarity 87.3%; Pred. No. 4.2e-44;  
Matches 103; Conservative 3; Mismatches 4; Indels 8; Gaps 1;  
QY 6 ESGGGLVQPKGSLKLSCAASGFTFNAYAMNVRQAPGKLEWVARIRKSNNTATFYADS 65  
DB 1 ESGGGLVQPKGSLKLSCAASGFTFNAYAMNVRQAPGKLEWVARIRKSNNTATFYADS 60  
QY 66 VKDRFTISRDDSQSMYLMQNNLKTEDTAMYYCVRGASGIDYMDYWGOGTSLTVSS 123  
DB 61 VKDRFTISRDDSQSMYLMQNNLKTEDTAMYYCVRGASGIDYMDYWGOGTSLTVSSA 110

RESULT 6  
AAU09919  
ID AAU09919 standard; Protein; 117 AA.

XX AC AAU09919;  
XX DT 18-JUN-2002 (first entry)  
XX DE Murine mAb 1D9 heavy chain variable region protein sequence.

XX KW Mouse; mAb 1D9 heavy chain variable region; vasotropic;  
KW antiinflammatory; collagen disease; immunosuppressive;  
KW antiasthmatic; insulin-dependent diabetes mellitus;  
KW inflammatory bowel disease; ulcerative colitis;  
KW graft rejection; allergic disease; antipsoriatic;  
KW antiarthritic; nephrotropic; antithyroid; restenosis;  
KW dermatological; anaphylaxis; cell adhesion inhibitor;  
KW vascular injury; autoimmune disease; immunoglobulin;  
KW complementarity determining region; CDR; CD18; CCR2;  
KW atherosclerosis.

XX OS Mus sp.

XX FH Key Location/Qualifiers  
FT Region 27..30  
FT /note= "Part of H1 structure loop"  
FT Region 31..35  
FT /note= "Complementarity determining region 1 (CDR1)"  
FT Region 50..68  
FT /note= "Complementarity determining region 2 (CDR2)"  
FT Region 101..106  
FT /note= "Complementarity determining region 3 (CDR3)"

XX WO200170266-A2.

XX 27-SEP-2001.

XX 15-MAR-2001; 2001WO-US08266.

XX 17-MAR-2000; 2000US-0528267.

XX (MILL-) MILLENNIUM PHARM INC.

XX Horvath CJ, Rao PE;

XX WPI; 2001-607511/69.

XX

PT Inhibiting stenosis or restenosis of a blood vessel following vascular  
PT injury or angioplasty in a subject by administering agent which  
PT inhibits recruitment or adhesion of neutrophils, mononuclear cells to  
PT injury site

XX Claim 30; Fig 16; 108pp; English.

XX CC The present invention relates to a new method of inhibiting stenosis or  
CC restenosis of a blood vessel following vascular injury in a subject. The  
CC new method comprises administering to the subject agents which inhibit  
CC the adhesion and/or recruitment of neutrophils and mononuclear cells to  
CC a site of vascular injury by binding CD18 or CCR2. The method of the  
CC invention inhibits stenosis or restenosis of a blood vessel following  
CC vascular injury arising from a vascular intervention procedure such as  
CC vascular by-pass or transplantation surgery. The method is also useful  
CC for treating a subject having an inflammatory disease or condition  
CC mediated by neutrophil and mononuclear cell activity e.g. asthma and  
CC graft versus host disease. Chronic inflammatory diseases of the lung,  
CC collagen diseases, and insulin-dependent diabetes mellitus can also be  
CC treated. The method is further useful for treating inflammatory bowel  
CC diseases, such as ulcerative colitis. Additional diseases or conditions  
CC include inflammatory or allergic diseases and conditions, including  
CC systemic anaphylaxis of hypersensitivity responses, drug allergies,  
CC psoriasis and inflammatory dermatoses, autoimmune diseases such as  
CC arthritis, graft rejection and other diseases including atherosclerosis.  
CC The present sequence represents the protein sequence of the mouse mAb  
CC 1D9 heavy chain variable region. This sequence contains three CDR's  
CC that can be used to inhibit stenosis or restenosis.

XX SQ Sequence 117 AA;

Query Match 80.3%; Score 519; DB 22; Length 117;  
Best Local Similarity 80.5%; Pred. No. 8.2e-43;  
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;

QY 1 EVMVESGGGLVQPKGSLKLSCAASGFTFNAYAMNVRQAPGKLEWVARIRKSNNTAT 60

DB 1 EVQLVESGGGLVQPKGSLKLSCAASGFTFNAYAMNVRQAPGKLEWVARIRKSNNTAT 60

QY 61 FYADSVKDRFTISRDDSQSMYLMQNNLKTEDTAMYYCVRGASGIDYMDYWGOGTSLT 120

DB 61 IYADSVKDRFTISRDDSQSMYLMQNNLKTEDTAMYYCVRGASGIDYMDYWGOGTSLT 114

QY 121 VSS 123

DB 115 VSS 117

RESULT 7  
AAE06947  
ID AAE06947 standard; Protein; 117 AA.

XX AC AAE06947;

XX DT 16-OCT-2001 (first entry)

XX DE Murine 1D9 antibody heavy chain variable (VH) region.

XX KW Murine; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;  
KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
KW fibrotic disease; angiolipolasty; acquired immune deficiency syndrome;  
KW AIDS; inflammatory glomerulopathy; vascular intervention; 1D9 antibody;  
KW neointimal hyperplasia; VH; heavy chain variable region.

XX OS Mus sp.

XX FH Key Location/Qualifiers  
FT Region 31..35  
FT /label= Complementarity\_determining\_region

```

FT Region 50..68
FT /label= Complementarity_determining_region
FT Region 101..106
FT /label= Complementarity_determining_region
XX WO200157226-A1.
XX 09-AUG-2001.
XX 02-FEB-2001; 2001WO-US03537.
XX 03-FEB-2000; 2000US-0497625.
XX (MILL-) MILLENNIUM PHARM INC.
XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;
XX WPI; 2001-488888/53.
XX Humanized immunoglobulin for treating a CC-chemokine receptor
XX 2-mediated disorder in a patient, comprises a binding specificity for
XX CCR2, and a non-human antigen binding region and human immunoglobulin
XX
XX Claim 8; Fig 8; 183pp; English.
XX The patent discloses a humanised antibody or its antigen-binding
XX fragment, having binding specificity for CC-chemokine receptor 2
XX (CCR2), comprising an antigen binding region of non-human origin
XX and at least a portion of an immunoglobulin of human origin. The
XX humanised antibodies are useful for inhibiting the interaction of
XX a cell expressing CCR2. They are useful for inhibiting or treating
XX HIV infection. The proteins of the invention are useful for inhibiting
XX leukocyte trafficking, for treating CCR2-mediated disorders such as
XX inflammatory disorder, autoimmune disorders such as rheumatoid
XX arthritis and multiple sclerosis, atherogenesis and atherosclerosis,
XX and for inhibiting restenosis. They are useful in therapy or diagnosis,
XX and in the manufacture of a medicament for treating CCR-2 mediated
XX disease. They are also useful for treating allergy, anaphylaxis,
XX malignancy, chronic and acute inflammation, histamine and IgE-
XX mediated allergic reaction, shock, stenosis, allograft rejection,
XX fibrotic disease, asthma, inflammatory glomerulopathies, acquired
XX immune deficiency syndrome (AIDS), restenosis associated with vascular
XX intervention, including angioplasty and/or stent placement in a mammal.
XX Humanised antibodies are also useful for inhibiting narrowing of the
XX lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of
XX a vessel in a mammal, preferably associated with vascular intervention.
XX The present sequence is the heavy chain variable (VH) region of
XX murine ID9 antibody.
XX Sequence 117 AA:
Query Match 80.3%; Score 519; DB 22; Length 117;
Best Local Similarity 80.5%; Pred. No. 8.2e-43;
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;
QY 1 EYVLVESGGGLVQPGKSLKLSAASGTFNNYAMNWRQAPGKGLEWARIKSNVAT 60
II IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Db 1 EYVLVESGGGLVQPGKSLKLSAASGTFNNYAMNWRQAPGKGLEWARIKSNVAT 60
II IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
QY 61 FYADSVKDRFTISRDDSQSMYLYQNNLTKETDAMYCYVRRGASGIDYAMDYWGQGTSLT 120
:IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Db 61 YVADSVKDRYTIISRDSESMFLQNNLTKETDAMYCYVTFYGVNGV-----WGTVTVT 114
:IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
QY 121 VSS 123
III
Db 115 VSS 117

RESULT 8
AAE07033
ID AAE07033 standard; Protein; 148 AA.
XX

```

```

AC AAE07033;
XX 16-OCT-2001 (first entry)
DT XX
DE XX
EX XX
KW Murine antibody ID9 heavy chain variable region.
KW Murine; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;
KW neuroprotective; immunosuppressive; human immunodeficiency virus;
KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;
KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;
KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;
KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;
KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;
KW AIDS; inflammatory glomerulopathy; vascular intervention;
KW neointimal hyperplasia; antibody ID9 heavy chain variable region.
XX Mus sp.
OS XX
PH Key
FT Peptide 1..19
FT /label= Signal_peptide
FT Protein 20..148
FT /note= "Murine mature antibody ID9 kappa
FT light chain variable region"
FT Region 20..137
FT /label= variable_region
FT Region 138..148
FT /label= constant_region
FT Misc-difference 148
FT /note= "Encoded by Gt"
PN WO200157226-A1.
XX 09-AUG-2001.
XX 02-FEB-2001; 2001WO-US03537.
XX 03-FEB-2000; 2000US-0497625.
XX (MILL-) MILLENNIUM PHARM INC.
XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;
XX WPI; 2001-488888/53.
XX N-PSDB; AAD13178.
XX Humanized immunoglobulin for treating a CC-chemokine receptor
XX 2-mediated disorder in a patient, comprises a binding specificity for
XX CCR2, and a non-human antigen binding region and human immunoglobulin
XX
XX Disclosure; Fig 21; 183pp; English.
XX The patent discloses a humanised antibody or its antigen-binding
XX fragment, having binding specificity for CC-chemokine receptor 2
XX (CCR2), comprising an antigen binding region of non-human origin
XX and at least a portion of an immunoglobulin of human origin. The
XX humanised antibodies are useful for inhibiting the interaction of
XX a cell expressing CCR2. They are useful for inhibiting or treating
XX HIV infection. The proteins of the invention are useful for inhibiting
XX leukocyte trafficking, for treating CCR2-mediated disorders such as
XX inflammatory disorder, autoimmune disorders such as rheumatoid
XX arthritis and multiple sclerosis, atherogenesis and atherosclerosis,
XX and for inhibiting restenosis. They are useful in therapy or diagnosis,
XX and in the manufacture of a medicament for treating CCR-2 mediated
XX disease. They are also useful for treating allergy, anaphylaxis,
XX malignancy, chronic and acute inflammation, histamine and IgE-
XX mediated allergic reaction, shock, stenosis, allograft rejection,
XX fibrotic disease, asthma, inflammatory glomerulopathies, acquired
XX immune deficiency syndrome (AIDS), restenosis associated with vascular
XX intervention, including angioplasty and/or stent placement in a mammal.
XX Humanised antibodies are also useful for inhibiting narrowing of the
XX lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of
XX a vessel in a mammal, preferably associated with vascular intervention.
XX The present sequence is the heavy chain variable (VH) region of
XX murine ID9 antibody.
XX Sequence 117 AA:
PS Disclosure; Fig 21; 183pp; English.
XX The patent discloses a humanised antibody or its antigen-binding
XX fragment, having binding specificity for CC-chemokine receptor 2
XX (CCR2), comprising an antigen binding region of non-human origin
XX and at least a portion of an immunoglobulin of human origin. The
XX humanised antibodies are useful for inhibiting the interaction of
XX a cell expressing CCR2. They are useful for inhibiting or treating
XX HIV infection. The proteins of the invention are useful for inhibiting
XX leukocyte trafficking, for treating CCR2-mediated disorders such as
XX inflammatory disorder, autoimmune disorders such as rheumatoid
XX arthritis and multiple sclerosis, atherogenesis and atherosclerosis,
XX and for inhibiting restenosis. They are useful in therapy or diagnosis,
XX and in the manufacture of a medicament for treating CCR-2 mediated
XX disease. They are also useful for treating allergy, anaphylaxis,
XX malignancy, chronic and acute inflammation, histamine and IgE-
XX mediated allergic reaction, shock, stenosis, allograft rejection,
XX fibrotic disease, asthma, inflammatory glomerulopathies, acquired
XX immune deficiency syndrome (AIDS), restenosis associated with vascular
XX intervention, including angioplasty and/or stent placement in a mammal.
XX Humanised antibodies are also useful for inhibiting narrowing of the
XX lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of
XX a vessel in a mammal, preferably associated with vascular intervention.
XX The present sequence is the heavy chain variable (VH) region of
XX murine ID9 antibody.
XX Sequence 117 AA:

```

CC a vessel in a mammal, preferably associated with vascular intervention.  
CC The present sequence is murine antibody ID9 heavy chain variable region.  
XX  
SQ Sequence 148 AA;  
Query Match 80.3%; Score 519; DB 22; Length 148;  
Best Local Similarity 80.5%; Pred. No. 1.1e-42;  
Matches 99; Conservative 11; Mismatches 7; Indels 6; Gaps 1;  
QY 1 EVMLVSGGLVQPKGSLKLSAASGFTFNAYAMNVRQAPGKLEWVARIRKSNYYAT 60  
DB 20 EVLVESGGGLVQPKGSLKLSAASGFTFNAYAMNVRQAPGKLEWVARIRKSNYYAT 79  
QY 61 FYADSVKDRFTISRDSQSMLYLQMNLLKTDYMYCVRGASGIDYAMDYWGQGTSLT 120  
DB 80 YYADSVKDRFTISRDSQSMLYLQMNLLKTDYMYCVRGASGIDYAMDYWGQGTSLT 133  
QY 121 VSS 123  
DB 134 VSS 136  
RESULT 9  
AAB10022  
ID AAB10022 standard; Protein; 125 AA.  
XX  
AC AAB10022;  
DT 01-NOV-2000 (first entry)  
XX  
DE H. pylori 26 kDa protein-binding antibody heavy chain protein.  
XX  
KW Acid-resistant microorganism; detection; faecal; intestine; infection;  
KW monoclonal antibody; heavy chain.  
XX  
OS Unidentified.  
XX  
PN WO200026671-A1.  
XX  
PD 11-MAY-2000.  
XX  
PF 29-OCT-1999; 99WO-EF08212.  
XX  
PR 29-OCT-1998; 98EP-0120517.  
PR 06-NOV-1998; 98EP-0120687.  
XX  
PA (CONN-) CONNEX GMBH.  
XX  
PI Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;  
PI Ringels A;  
XX  
WPI; 2000-365747/31.  
DR N-PSDB; AAA40202.  
XX  
XX Detecting infection by acid-fast microbes for diagnosis of Helicobacter  
PT pylori, comprises reacting a faecal sample with two binding reagents for  
PT antigens that survive intestinal passage -  
XX  
PS Disclosure; Fig 4; 84pp; German.  
XX  
CC This invention describes a novel method for the detection of a mammalian  
CC infection by an acid-resistant microorganism (A) by treating a faecal  
CC sample with at least two different monoclonal antibodies (mAb) (or their  
CC fragments or derivatives) or aptamers (collectively (I)) and detecting  
CC formation of a complex (C) between (I) and the corresponding antigen of  
CC (A). The first and second (I) bind to epitopes of different antigens  
CC (Ag). These epitopes are present, after passage through the intestines,  
CC in at least some mammals, and have either: (i) their native structure;  
CC or (ii) a structure against which an antibody is produced by an animal  
CC infected or immunized with (A), or its extract, lysate, derived protein  
CC or fragment, or with a synthetic peptide. Practically all mammals display  
CC at least one of the specified epitopes. The method is used to detect  
CC infection by acid-fast bacteria, particularly of the genera Helicobacter,

CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,  
CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used  
CC therapeutically. The method is direct and non-invasive, and provides an  
CC inexpensive and easily standardizable diagnosis, despite possible  
CC degradation of antigens during passage through the intestines. This  
CC sequence represents the H. pylori 26 kDa protein-binding antibody  
CC (DMS ACC2335) heavy chain which is used to illustrate the method of the  
CC invention.  
XX  
SQ Sequence 125 AA;  
Query Match 79.6%; Score 514.5; DB 21; Length 125;  
Best Local Similarity 82.0%; Pred. No. 2.4e-42;  
Matches 100; Conservative 8; Mismatches 13; Indels 1; Gaps 1;  
QY 3 MLVESGGGLVQPKGSLKLSAASGFTFNAYAMNVRQAPGKLEWVARIRKSNYYAT 62  
DB 4 LLEESGGGLVQPKGSLKLSAASGFTFNAYAMNVRQAPGKLEWVARIRKSNYYAT 63  
QY 63 ADSVKDRFTISRDSQSMLYLQMNLLKTDYMYCVR-RGASGIDYAMDYWGQGTSLTV 121  
DB 64 ANSVKDRFTISRDSQSMLYLQMNLLKTDYMYCVRDHDKFPFYALDYWGPGTLTV 123  
QY 122 SS 123  
DB 124 SS 125  
RESULT 10  
AAO14978  
ID AAO14978 standard; protein; 117 AA.  
XX  
AC AAO14978;  
DT 05-SEP-2002 (first entry)  
XX  
DE Mouse heavy chain variable region (mAb 1D9).  
XX  
KW Mouse; graft rejection; CC chemokine receptor 2 antagonist;  
KW CCR2 antagonist; anti-CCR2 antibody; kidney transplant; liver transplant;  
KW lung transplant; heart-lung transplant; pancreas transplant;  
KW bowel transplant; heart transplant; graft versus host disease;  
KW chronic graft rejection; antibody heavy chain variable region; mAb 1D9.  
XX  
OS Mus musculus.  
XX  
PN US2002042370-A1.  
XX  
PD 11-APR-2002.  
XX  
PF 13-APR-2001; 2001US-0835087.  
XX  
PR 14-APR-2000; 2000US-0549448.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Hancock WW;  
XX  
XX WPI; 2002-351365/38.  
XX  
PT Inhibiting graft rejection, graft versus host disease or chronic  
PT rejection of a transplanted graft, involves administering a CCR2  
PT antagonist -  
XX  
PS Disclosure; Fig 2; 16pp; English.  
XX  
CC The invention comprises a method of inhibiting graft rejection, graft  
CC versus host disease or chronic rejection of a transplanted graft. The  
CC method involves administering an antagonist of CC chemokine receptor 2  
CC (CCR2) and optionally an immunosuppressive agent. The CCR2 antagonist may  
CC be an anti-CCR2 antibody (i.e. containing light and heavy chain  
CC complementarity determining regions from various non-human origins). CCR2  
CC is known to be involved in the rejection of transplanted grafts. The



CC method of the invention is useful for inhibiting graft rejection -  
CC particularly allografts such as kidney, liver, lung, heart-lung,  
CC pancreas, bowel and heart. The method of the invention is also useful for  
CC inhibiting graft versus host disease and for inhibiting chronic rejection  
CC of a transplanted graft. The present amino acid sequence represents a  
CC mouse antibody heavy chain variable region (mAb 1D9).  
XX  
SQ

Sequence 117 AA;

Query Match 78.9%; Score 510; DB 23; Length 117;

Best Local Similarity 79.7%; Pred. No. 6.1e-42;

Matches 98; Conservative 11; Mismatches 8; Indels 6; Gaps 1;

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 61 FYADSVKDRFTISRDSDQSMYLMQNNLKTEDTAMYCYVRRGASGIDYAMDYWGQTSLT 120

DB 61 FYADSVKDRFTISRDSDQSMYLMQNNLKTEDTAMYCYVRRGASGIDYAMDYWGQTSLT 120

QY 121 VSS 123

DB 115 VSS 117

RESULT 11

ID AAU09930 standard; Protein; 117 AA.

XX AAU09930;

XX 18-JUN-2002 (first entry)

XX Humanised 1D9 heavy chain variable region, 1D9RH protein sequence.

XX Human; mouse; 4B4'CL heavy chain variable region; vasotropic;

XX antiinflammatory; collagen disease; immunosuppressive;

KW antiasthmatic; insulin-dependent diabetes mellitus;

KW inflammatory bowel disease; ulcerative colitis; 1D9RH;

KW graft rejection; allergic disease; antipsoriatic;

KW antiarthritic; nephrotropic; antithyroid; restenosis;

KW dermatological; anaphylaxis; cell adhesion inhibitor;

KW vascular injury; autoimmune disease; immunoglobulin;

KW complementarity determining region; CDR; CD18; CCR2;

XX atherosclerosis; mutant; mucin.

OS Chimeric - Homo sapiens.

OS Chimeric - Mus sp.

OS Chimeric - Synthetic.

FH Key Location/Qualifiers

FT Region 27..30

FT MISC-difference 28

FT /note- "Part of H1 structure loop"

FT /note- "Substitution of Thr residue normally present

FT in human 4B4'CL sequence (AAU09926) by Ser

FT residue normally present in mouse mAb 1D9 heavy

FT chain sequence (AAU09919)."

FT MISC-difference 30

FT /note- "Substitution of Ser residue normally present

FT in human 4B4'CL sequence (AAU09926) by Asn

FT residue normally present in mouse mAb 1D9 heavy  
FT chain sequence (AAU09919)."  
FT Region 50..68  
FT /note- "Complementarity determining region 2 (CDR2),  
FT grafted from mouse mAb 1D9 heavy chain  
FT sequence (AAU09919)."

FT MISC-difference 70

FT /note- "Substitution of Phe residue normally present  
FT in human 4B4'CL sequence (AAU09926) by Tyr

FT residue normally present in mouse mAb 1D9 heavy

FT chain sequence (AAU09919)."

FT MISC-difference 99

FT /note- "Substitution of Thr residue normally present

FT in human 4B4'CL sequence (AAU09926) by Val

FT residue normally present in mouse mAb 1D9 heavy

FT chain sequence (AAU09919)."

FT Region 101..108

FT /note- "Complementarity determining region 3 (CDR3),

FT grafted from mouse mAb 1D9 heavy chain

FT sequence (AAU09919)."

XX WO200170266-A2.

XX 27-SEP-2001.

XX 15-MAR-2001; 2001WO-US08266.

XX 17-MAR-2000; 2000US-0528267.

XX (MILL-) MILLENNIUM PHARM INC.

XX Horvath CJ, Rao PE;

XX WPI; 2001-607511/69.

XX Claim 32; Fig 18; 108pp; English.

XX The present invention relates to a new method of inhibiting stenosis or

XX restenosis of a blood vessel following vascular injury in a subject. The

XX new method comprises administering to the subject agents which inhibit

XX the adhesion and/or recruitment of neutrophils and mononuclear cells to

XX a site of vascular injury by binding CD18 or CCR2. The method of the

XX invention inhibits stenosis or restenosis of a blood vessel following

XX vascular injury arising from a vascular intervention procedure such as

XX vascular by-pass or transplantation surgery. The method is also useful

XX for treating a subject having an inflammatory disease or condition

XX mediated by neutrophil and mononuclear cell activity e.g. asthma and

XX graft versus host disease. Chronic inflammatory diseases of the lung,

XX collagen diseases, and insulin-dependent diabetes mellitus can also be

XX treated. The method is further useful for treating inflammatory bowel

XX diseases, such as ulcerative colitis. Additional diseases or conditions

XX include inflammatory or allergic diseases and conditions, including

XX systemic anaphylaxis of hypersensitivity responses, drug allergies,

XX psoriasis and inflammatory dermatoses, autoimmune diseases such as

XX arthritis, graft rejection and other diseases including atherosclerosis.

CC This sequence represents the variable region of one of several humanised  
CC 1D9 heavy chains (AAU09927-AAU09930). These heavy chains were used in  
CC the invention for the production of anti-CCR2 antibody or antigen-binding  
CC fragment.  
XX  
SQ

Sequence 117 AA;

Query Match 76.3%; Score 493; DB 22; Length 117;

Best Local Similarity 76.4%; Pred. No. 2.7e-40;

Matches 94; Conservative 14; Mismatches 9; Indels 6; Gaps 1;

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

QY 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60

DB 1 EYMLVSGGLVOPKSGLSKLSAASGFTFNVMNVRAPGKGLWVARIRSKSNYAT 60







```

Db      1  EVLVESGGGLVPGGSLRLSCAAGSFENYAMNVRQAPGKGLIEWARIETKNNNTAT  60
Qy      61  FYADSVKDRFTISRDDSSQSMLYLQMNNLKTDETAMTYCYCVRGASGIDYAMDYWGQGTSLT  120
       :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61  YYADSVKDRYTIISRDDSKNTLYLQMNLSLKTEDTAVYCYCTTFYGVNGV-----WGQGTLLVT  114

Qy      121  VSS 123
       |||
Db      115  VSS 117

```

Search completed: November 27, 2002, 07:17:53  
Job time : 70.9112 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OW protein - protein search, using sw model

Run on: November 27, 2002, 07:15:05 ; Search time 33.7181 Seconds  
(without alignments)  
350.688 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 646

Sequence: 1 EVMLVSGGLVQPKGSLK.....SGIDYAMYWGQGTSLTVSS 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	528.5	81.8	141	2 I32513	Ig heavy chain pre
2	526.5	81.5	137	2 S42467	Ig heavy chain V r
3	522	80.8	121	2 A41940	Ig heavy chain V r
4	494.5	76.5	110	2 PH1091	Ig heavy chain V r
5	489	75.7	119	2 A27630	Ig heavy chain pre
6	480.5	74.4	110	2 PH1092	Ig heavy chain V r
7	477	73.8	110	2 PH1093	Ig heavy chain V r
8	467.5	72.4	101	2 B42575	Ig heavy chain V r
9	463	71.7	127	2 S58213	Ig heavy chain V r
10	456	70.5	100	2 S26452	Ig heavy chain V r
11	450	69.7	123	2 S32186	Ig heavy chain V r
12	449.5	69.6	124	2 S20775	Ig heavy chain V r
13	446	69.0	115	1 AVMS06	Ig heavy chain V-I
14	444	68.7	143	2 S3624	Ig heavy chain V r
15	443.5	68.7	115	2 S38714	Ig heavy chain V r
16	442.5	68.5	122	2 S20642	Ig heavy chain V r
17	441.5	68.3	118	2 A31485	Ig heavy chain V r
18	441	68.3	119	2 D30540	Ig heavy chain V r
19	439.5	68.0	111	1 MHMS76	Ig heavy chain V-I
20	439	68.0	119	2 S31107	Ig heavy chain V-I
21	439	68.0	121	2 H36005	Ig heavy chain - h
22	439	68.0	139	2 S31678	Ig heavy chain V r
23	438.5	67.9	126	2 S44107	Ig heavy chain V r
24	438.5	67.9	140	2 S70442	Ig heavy chain V-D
25	438	67.8	113	1 AVMSB7	Ig heavy chain pre
26	438	67.8	121	2 S31106	Ig heavy chain V-I
27	438	67.8	124	2 PT0388	Ig heavy chain - h
28	437	67.6	115	2 A25803	Ig heavy chain V r
29	436	67.5	113	1 AVMSAB	Ig heavy chain V-I

30	436	67.5	115	1 AVMS82	Ig heavy chain V-I
31	436	67.5	117	2 S31109	Ig heavy chain - h
32	435	67.3	160	2 S05271	Ig heavy chain pre
33	433	67.0	140	2 S31588	Ig heavy chain V r
34	432	66.9	123	2 A36006	Ig heavy chain V r
35	431.5	66.8	121	2 I27887	Ig heavy chain V r
36	431	66.7	118	2 S20641	Ig heavy chain V r
37	431	66.7	122	2 E27888	Ig heavy chain V r
38	431	66.7	123	2 S26794	Ig heavy chain V r
39	431	66.7	125	2 S30531	Ig heavy chain V r
40	431	66.7	140	2 S31686	Ig heavy chain V r
41	430.5	66.6	119	2 F27888	Ig heavy chain V r
42	430.5	66.6	120	2 S44111	Ig heavy chain V-D
43	430.5	66.6	147	2 I37780	Ig variable region
44	430	66.6	113	1 AVMS61	Ig heavy chain V-I
45	430	66.6	119	2 C36005	Ig heavy chain V r

ALIGNMENTS

RESULT 1

I32513

Ig heavy chain precursor V region (MRL4) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-May-1990 #sequence\_revision 31-Dec-1990 #text\_change 23-Jul-1999

C:Accession: I32513

R:Kofler, R.; Strohal, R.; Balderas, R.S.; Johnson, M.E.; Noonan, D.J.; Duchosal, M.A

J. Clin. Invest. 82, 852-860, 1988

A:Title: Immunoglobulin kappa light chain variable region gene complex organization a

A:Reference number: A94689; MUID:88331394; PMID:3138286

A:Accession: I32513

A:Molecule type: DNA

A:Residues: 1-141 <KOF>

A:Cross-references: GB:M20829; NID:g196951; PIDN:AAA38849.1; PID:g196952

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:34-119/Domain: immunoglobulin homology <IMM>

Query Match 81.8%; Score 528.5; DB 2; Length 141;

Best Local Similarity 83.7%; Pred. No. 9.5e-42;

Matches 103; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

Qy 1 EVMLVSGGLVQPKGSLKSCAASGFTFNYYAMNWRQAPGKGLWVARRSKNYYAT 60

Db 20 EVQLVETGGGLVQPKGSLKSCPASGFSFNTNMMWRQAPGKGLWVARRSKNYYAT 79

Qy 61 FYADSVKDRFTISRDDSQSMLYLQNNLKTEDTAMYYCVRRGASGIDYAMDWGGGTSLT 120

Db 80 YYADSVKDRFTISRDDSQSMLYLQNNLKTEDTAMYYCVRDAANWSAN-FAYWGGGTSLT 138

Qy 121 VSS 123

Db 139 VSA 141

RESULT 2

S42467

Ig heavy chain V region precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999

C:Accession: S42467

R:Shiyanov, P.A.; Bespalov, I.A.; Terletskaya, H.N.; Deyev, S.M.

submitted to the EMBL Data Library, March 1994

A:Reference number: S42466

A:Accession: S42467

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-137 <SHI>

A:Cross-references: EMBL:X78107; NID:g460798; PIDN:CAAS4997.1; PID:g460799

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:34-119/Domain: immunoglobulin homology <IMM>





C;Keywords: heterotetramer; immunoglobulin  
F;15-100/Domain: immunoglobulin homology <IMM>

Query Match 74.4%; Score 480.5; DB 2; Length 110;  
Best Local Similarity 82.6%; Pred. No. 1.9e-37;  
Matches 95; Conservative 3; Mismatches 10; Indels 7; Gaps 2;

Qy 1 EVMLVESGGGLVQPKGSLKLSAASGFTFNMYAMNVRQAPCKGLEWVARIRKSNYYAT 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 EVQLVETGGGLVQPKGSLKLSAASGTSFTNMYAMNVRQAPCKGLEWVARIRKSNYYAT 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Qy 61 FYADSVKDRFTISRDDSOSMLYLQNNLKTEDTAMYYCVRRGASGIDYA--MDYW 113  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 YYADSVKDRFTISRDDSOSMLYLQNNLKIEDTAMYYCVRD-----DYVAFAYW 110  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 7

PH1093

Ig heavy chain V region (clone SL7.161) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C;Accession: PH1093

R;Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A;Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B c

A;Reference number: PH0971; MUID:92381444; PMID:1512540

A;Accession: PH1093

A;Status: nucleic acid sequence not shown.

A:Molecule type: mRNA

A;Residues: 1-110 &lt;TIL&gt;

A;Experimental source: B cell, strain [NZB x NZW]F1

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 73.8%; Score 477; DB 2; Length 110;  
Best Local Similarity 90.0%; Pred. No. 4.1e-37;  
Matches 90; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVMLVESGGGLVQPKGSLKLSAASGFTFNMYAMNVRQAPCKGLEWVARIRKSNYYAT 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 EVQLVETGGGLVQPKGSLKLSAASGTSFTNMYAMNVRQAPCKGLEWVARIRKSNYYAT 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 FYADSVKDRFTISRDDSOSMLYLQNNLKTEDTAMYYCVR 100

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 61 YYADSVKDRFTISRDDSOSMLYLQNNLKTDDETAMYYCVR 100

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 8

B42575

Ig heavy chain V region (anti-angiotensin II) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 21-Jan-2000

C;Accession: B42575

R;Garcia, K.C.; Desiderio, S.V.; Ronco, P.M.; Verroust, P.J.; Amzel, L.M.

Science 257, 528-531, 1992

A;Title: Recognition of angiotensin II: antibodies at different levels of an idiotypic r

A;Reference number: A42575; MUID:92342952; PMID:1636087

A;Accession: B42575

A:Molecule type: mRNA

A;Residues: 1-101 &lt;GAR&gt;

A;Cross-references: GB:540679; NID:9252098; PIDN:AAB22669.1; PID:g252099

A;Experimental source: BALB/c mice, myeloma cell line NS-1

A;Note: sequence extracted from NCBI backbone (NCBIN:109337, NCBIIP:109338)

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: immunoglobulin

F;7-92/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 72.4%; Score 467.5; DB 2; Length 101;  
Best Local Similarity 87.6%; Pred. No. 2.8e-36;  
Matches 92; Conservative 2; Mismatches 6; Indels 5; Gaps 2;

Qy 9 GGLVQPKGSLKLSAASGFTFNMYAMNVRQAPCKGLEWVARIRKSNYYATFYADSVKD 68

Db 1 GGLVQPKGSLKLSAASGFTFNMYAMNVRQAPCKGLEWVARIRKSGTYTYTADSVKD 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Qy 69 RFTISRDDSOSMLYLQNNLKTEDTAMYYCVRRGASGID-YAMDY 112  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 RFTISRDDSOSMLYLQNNLKTEDTAMYYCVR---GYDGYAMDY 101  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 9

S58213

Ig heavy chain V region (anti-F(ab')2) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1996 #sequence\_revision 12-Apr-1996 #text\_change 23-Jul-1999

C;Accession: S58213; S58212

R;Welschhof, M.; Terness, P.; Stanescu, D.; Zewe, M.; Hain, C.H.; Doebel, S.; Breitlin

submitted to the EMBL Data Library, July 1995

A;Description: Characterization of heavy and light chain immunoglobulin variable regi

A;Reference number: S58206

A;Accession: S58213

A:Molecule type: mRNA

A;Residues: 1-127 &lt;WEL&gt;

A;Cross-references: EMBL:X89055; NID:g929638; PIDN:CAA61442.1; PID:g929639; EMBL:X890

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 71.7%; Score 463; DB 2; Length 127;

Best Local Similarity 71.9%; Pred. No. 9.2e-36;

Matches 92; Conservative 11; Mismatches 19; Indels 6; Gaps 2;

Qy 1 EVMLVESGGGLVQPKGSLKLSAASGFTFNMYAMNVRQAPCKGLEWVARIRKSNYYAT 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 EVQLVESGGGLVQPKGSLKLSAASGFTTSGTHHWVRQASGKGLEWVGRIKNDNSYAT 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 FYADSVKDRFTISRDDSOSMLYLQNNLKTEDTAMYYCVR-----RGASGIDYAMDYWGQ 115

|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 61 AYAASVKGRTISRDDSSENTAYLQNNSLKIEDTAVYYCTRGSSMVRGVNGY-YGMDVWGQ 119

|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 116 GTSLTVSS 123

Db 120 GTTVTVSS 127

## RESULT 10

S26462

Ig heavy chain V region - mouse

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997

C;Accession: S26462

R;Kavaler, J.

submitted to the EMBL Data Library, April 1991

A;Reference number: S26459

A;Accession: S26462

A;Status: preliminary

A:Molecule type: mRNA

A;Residues: 1-100 &lt;KAV&gt;

A;Cross-references: EMBL:X59104; NID:g51920; PID:g51921

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;7-92/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 70.6%; Score 456; DB 2; Length 100;

Best Local Similarity 84.3%; Pred. No. 3.2e-35;

Matches 91; Conservative 4; Mismatches 3; Indels 10; Gaps 3;

Qy 11 LVQPKGS--LKLSAASGFTFNMYAMNVRQAPCKGLEWVARIRKSNYYATFYADSVKD 68  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 1 LVQPKGS--LKLSAASGFTFNMYAMNVRQAPCKGLEWVARIRKSNYYATFYADSVKD 60

|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 69 RFTISRDDSOSMLYLQNNLKTEDTAMYYCVRRGASGIDYAMDYWGQ 116

|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 61 RFTISRDDSOSMLYLQNNLKTEDTAMYYCVS-----DFS--HWGQ 100

|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 11

S32186  
Ig heavy chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 23-Jul-1999  
C:Accession: S32186  
R:Izui, S.  
submitted to the EMBL Data Library, February 1993  
A:Reference number: S32185  
A:Accession: S32186  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-123 <I2>  
A:Cross-references: EMBL:X70093; NID:g288249; PIDN:CAA49698.1; PID:g288250  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>

Query Match 69.7%; Score 450; DB 2; Length 123;  
Best Local Similarity 69.9%; Pred. No. 1.4e-34;  
Matches 86; Conservative 12; Mismatches 25; Indels 0; Gaps 0;  
QY 1 EYVLVESGGGLVQPGKSLKSCAASGFTFNYYAMNVRQAPGKLEWVARIRSKSNYYAT 60  
DB 1 EVKLVESGGGLVQPGKSLKSCATSGFTFDYYMSWVRQPPGKALEWLGFIIRKANGYTT 60  
QY 61 FYADSVKDRFTISRDDSQSMVLYQMNNLKTEDTAMYYCYVRRGASGIDYAMDYWGQGTSLT 120  
DB 61 EYASVKGKFTISRDNSSQILYQMNNLRAEDSATYICARSYYGSSRYAMDYWGQGTSVT 120  
QY 121 VSS 123  
DB 121 VSS 123

## RESULT 12

S20775  
Ig heavy chain V region - human  
C:Species: Homo sapiens (man)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C:Accession: S20775  
R:Morcarl, F.; Wang, J.; Schroeder, H.W.  
submitted to the EMBL Data Library, April 1992  
A:Description: Analysis of human cord blood Ig heavy chain IgA and IgG repertoire.  
A:Reference number: S20764  
A:Accession: S20775  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-124 <MOR>  
A:Cross-references: EMBL:Z11950; NID:g33882; PIDN:CAA78007.1; PID:g33883  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>

Query Match 69.6%; Score 449.5; DB 2; Length 124;  
Best Local Similarity 71.0%; Pred. No. 1.6e-34;  
Matches 88; Conservative 14; Mismatches 21; Indels 1; Gaps 1;  
QY 1 EYVLVESGGGLVQPGKSLKSCAASGFTFNYYAMNVRQAPGKLEWVARIRSKSNYYAT 60  
DB 1 EYVLVESGGGLVQPGKSLKSCAASGFTFSDYYMDVRQAPAKGLEWLTARIRKANSYTT 60  
QY 61 FYADSVKDRFTISRDDSQSMVLYQMNNLKTEDTAMYYCYVRRGASGID-YAMDYWGQGTSL 119  
DB 61 EYASVKGKFTISRDNSSMNSLSLQNNLKTEDTAIYYCVTRCHGMDVWGQGTTV 120  
QY 120 TVSS 123  
DB 121 TVSS 124

## RESULT 13

## AVMS06

Ig heavy chain V-III region (J606) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996  
C:Accession: C92811; A02072  
R:Johnson, N.; Slankard, J.; Paul, L.; Hood, L.  
J. Immunol. 128, 302-307, 1982  
A:Title: The complete V domain amino acid sequences of two myeloma inulin-binding pro  
A:Reference number: A92811; MUID:82099361; PMID:6798111  
A:Accession: C92811  
A:Molecule type: protein  
A:Residues: 1-115 <JOH>  
C:Comment: This chain was isolated from a myeloma protein that binds inulin.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>  
F:22-98/Disulfide bonds: #status predicted

Query Match 69.0%; Score 446; DB 1; Length 115;  
Best Local Similarity 70.7%; Pred. No. 3.1e-34;  
Matches 87; Conservative 10; Mismatches 18; Indels 8; Gaps 1;  
QY 1 EYVLVESGGGLVQPGKSLKSCAASGFTFNYYAMNVRQAPGKLEWVARIRSKSNYYAT 60  
DB 1 EVKLVESGGGLVQPGKSLKSCVAGSFTFSNYMMNVRQSPKLEWVAEIRLKSNNYAT 60  
QY 61 FYADSVKDRFTISRDDSQSMVLYQMNNLKTEDTAMYYCYVRRGASGIDYAMDYWGQGTSLT 120  
DB 61 HYAESVKGKFTISRDDSKSVLYQMNNLRAEDTGIIYCTGTGA-----YWGQGTSLT 112  
QY 121 VSS 123  
DB 113 VSA 115

## RESULT 14

S23624  
Ig heavy chain V region - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S23624  
R:Olee, T.; Lu, E.W.; Huang, D.F.; Soto-Gil, R.W.; Defios, M.; Kozin, F.; Carson, D.A.  
J. Exp. Med. 175, 831-842, 1992  
A:Title: Genetic analysis of self-associating immunoglobulin G rheumatoid factors fro  
A:Reference number: S23623; MUID:92156804; PMID:1740665  
A:Accession: S23624  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-143 <OLE>  
A:Cross-references: EMBL:X59703; NID:g32012; PIDN:CAA42224.1; PID:g32013  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 68.7%; Score 444; DB 2; Length 143;  
Best Local Similarity 70.7%; Pred. No. 5.9e-34;  
Matches 87; Conservative 14; Mismatches 16; Indels 6; Gaps 2;  
QY 1 EYVLVESGGGLVQPGKSLKSCAASGFTFNYYAMNVRQAPGKLEWVARIRSKSNYYAT 60  
DB 1 EYVLVESGGGLVQPGKSLKSCAASGFTFSNYSMMNVRQAPGKLEWVSTISSSS--TI 58  
QY 61 FYADSVKDRFTISRDDSQSMVLYQMNNLKTEDTAMYYCYVRRGASGIDYAMDYWGQGTSLT 120  
DB 59 YYADSVKGRFTISRDNKNSLYQMNNLRAEDTAVYYCARSYRG----GDYWGQGTSLT 114  
QY 121 VSS 123  
DB 115 VSS 117

## RESULT 15

S38714

Search completed: November 27, 2002, 07:24:15  
Job time : 35.7181 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	446	69.0	115	1	HV32_MOUSE	P01801 mus musculus
2	439.5	68.0	111	1	HV35_MOUSE	P01804 mus musculus
3	438	67.8	113	1	HV30_MOUSE	P01799 mus musculus
4	436	67.5	113	1	HV27_MOUSE	P01796 mus musculus
5	436	67.5	115	1	HV33_MOUSE	P01802 mus musculus
6	430	66.6	113	1	HV28_MOUSE	P01797 mus musculus
7	427	66.1	142	1	HV01_RAT	P01805 rattus norv
8	426	65.9	113	1	HV29_MOUSE	P01798 mus musculus
9	422	65.3	113	1	HV31_MOUSE	P01800 mus musculus
10	421	65.2	123	1	HV25_MOUSE	P01794 mus musculus
11	416.5	64.5	136	1	HV16_MOUSE	P01783 mus musculus
12	411.5	63.7	122	1	HV21_MOUSE	P01790 mus musculus
13	410.5	63.5	122	1	HV3G_HUMAN	P01768 homo sapien
14	407	63.0	123	1	HV18_MOUSE	P01787 mus musculus
15	406	62.8	113	1	HV34_MOUSE	P01803 mus musculus
16	406	62.8	123	1	HV19_MOUSE	P01788 mus musculus
17	405.5	62.8	122	1	HV20_MOUSE	P01789 mus musculus
18	405	62.7	123	1	HV22_MOUSE	P01791 mus musculus
19	403	62.4	144	1	HV26_MOUSE	P01795 mus musculus
20	399	61.8	121	1	HV3J_HUMAN	P01771 homo sapien
21	398	61.6	123	1	HV24_MOUSE	P01773 mus musculus
22	397	61.5	123	1	HV23_MOUSE	P01792 mus musculus
23	393.5	60.9	116	1	HV3T_HUMAN	P01781 homo sapien
24	389	60.2	117	1	HV17_MOUSE	P01786 mus musculus
25	388	60.1	120	1	HV3E_HUMAN	P01766 homo sapien
26	385.5	59.7	122	1	HV3H_HUMAN	P01769 homo sapien
27	385	59.6	115	1	HV3D_HUMAN	P01765 homo sapien
28	383.5	59.4	126	1	HV3K_HUMAN	P01772 homo sapien
29	378.5	58.6	119	1	HV38_MOUSE	P01808 mus musculus
30	377	58.4	117	1	HV42_MOUSE	P01812 mus musculus
31	377	58.4	119	1	HV41_MOUSE	P01770 homo sapien
32	376	58.2	117	1	HV3F_HUMAN	P01767 homo sapien
33	375	58.0	117	1	HV41_MOUSE	P01811 mus musculus

DE Ig heavy chain V-III region HPC76 (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81013937; PubMed=6251474;  
 RA Bernard O., Gough N.M.;  
 RT "Nucleotide sequence of immunoglobulin heavy chain joining segments  
 between translocated VH and mu constant regions genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 77:3630-3634 (1980).  
 CC -1- MISCELLANEOUS: THE SEQUENCE OF THE FIRST 197 RESIDUES OF THE C  
 REGION WAS ALSO DETERMINED & DIFFERS IN ONLY 3 POSITIONS FROM THE  
 CORRESPONDING PORTION OF THE MOUSE MOPC 104E MU CHAIN.  
 DR PIR: A02074; MMS76.  
 DR HSSP: P01789; IMCP.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig\_1.  
 DR SMART: SM00406; Igv; 1.  
 DR Immunoglobulin V region.  
 FT NON\_TER 1 111  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12304 MW; 0EDE98EC7348056A CRC64;  
 Query Match 68.0%; Score 439.5; DB 1; Length 111;  
 Best Local Similarity 72.9%; Pred. No. 3.1e-40;  
 Matches 86; Conservative 9; Mismatches 16; Indels 7; Gaps 2;  
 Qy 6 EGGGLVQPGKSLKSCAASGFTFNMYAMNVRQAPGKLEWVARIRSKSNYYAT 65  
 Db 1 EGGGLVQPGGSMKLSVASGFTFSNYMNMVRSQPEKGLWVAEIRLKS-HYAT 59  
 Qy 66 VKDFTISRDDQSOMLYLQNMNLTEDTAMYYCVRGASGIDYAMDYWGQGTSLTSS 123  
 Db 60 VKGRTISRDDSKSVYLQNMNLRADDTGIYCTRPVPP-----DYWGQGTSLTSS 111  
 RESULT 3  
 HV30\_MOUSE  
 ID HV30\_MOUSE STANDARD; PRT; 113 AA.  
 AC P01799;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region ABE-47N.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77134726; PubMed=402936;  
 RA Vrana M., Rudikoff S., Potter M.;  
 RT "Heavy-chain variable-region sequence from an inulin-binding myeloma  
 protein.";  
 RL Biochemistry 16:1170-1175 (1977).  
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
 BINDS INULIN.  
 DR PIR: A90400; AVMSB7.  
 DR HSSP: P01810; 2FBJ.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig\_1.  
 DR SMART: SM00406; Igv; 1.  
 DR Immunoglobulin V region.  
 FT DISULFID 22 98  
 FT NON\_TER 113 113  
 SQ SEQUENCE 113 AA; 12675 MW; 76638C16C779845E CRC64;  
 Query Match 67.8%; Score 438; DB 1; Length 113;  
 Best Local Similarity 70.2%; Pred. No. 4.5e-40;  
 Matches 85; Conservative 10; Mismatches 18; Indels 8; Gaps 1;  
 Qy 1 EVMLVESGGGLVQPGKSLKSCAASGFTFNMYAMNVRQAPGKLEWVARIRSKSNYYAT 60  
 Db 1 EVKLEESGGGLVQPGGSMKLSVASGFTFSNYMNMVRSQPEKGLWVAEIRLKS-HYAT 60  
 Qy 61 FYADSVKDRFTISRDDQSOMLYLQNMNLTEDTAMYYCVRGASGIDYAMDYWGQGTSLT 120  
 Db 61 HYAESVKGRTISRDDSKSVYLQNMNLRADDTGIYCTTGFA-----YWGQGTSLT 112  
 Qy 121 V 121  
 Db 113 V 113  
 RESULT 5  
 HV33\_MOUSE  
 ID HV33\_MOUSE STANDARD; PRT; 115 AA.  
 AC P01802;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region W3082.  
 OS Mus musculus (Mouse).  
 Query Match 67.8%; Score 438; DB 1; Length 113;  
 Best Local Similarity 70.2%; Pred. No. 4.5e-40;  
 Matches 85; Conservative 10; Mismatches 18; Indels 8; Gaps 1;

Matches 85; Conservative 10; Mismatches 18; Indels 8; Gaps 1;  
 Qy 1 EVMLVESGGGLVQPGKSLKSCAASGFTFNMYAMNVRQAPGKLEWVARIRSKSNYYAT 60  
 Db 1 EVKLEESGGGLVQPGGSMKLSVASGFTFSNYMNMVRSQPEKGLWVAEIRLKS-HYAT 60  
 Qy 61 FYADSVKDRFTISRDDQSOMLYLQNMNLTEDTAMYYCVRGASGIDYAMDYWGQGTSLT 120  
 Db 61 HYAESVKGRTISRDDSKSVYLQNMNLRADDTGIYCTTGFA-----YWGQGTSLT 112  
 Qy 121 V 121  
 Db 113 V 113  
 RESULT 4  
 HV27\_MOUSE  
 ID HV27\_MOUSE STANDARD; PRT; 113 AA.  
 AC P01796;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region A4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=78158406; PubMed=417344;  
 RA Vrana M., Rudikoff S., Potter M.;  
 RT "Sequence variation among heavy chains from inulin-binding myeloma  
 proteins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961 (1978).  
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
 BINDS INULIN.  
 DR PIR: A93818; AVMSAB.  
 DR HSSP: P01810; 2FBJ.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig\_1.  
 DR SMART: SM00406; Igv; 1.  
 DR Immunoglobulin V region.  
 FT DISULFID 22 98  
 FT NON\_TER 113 113  
 SQ SEQUENCE 113 AA; 12675 MW; 76658C121C598285 CRC64;  
 Query Match 67.5%; Score 436; DB 1; Length 113;  
 Best Local Similarity 70.2%; Pred. No. 7.4e-40;  
 Matches 85; Conservative 10; Mismatches 18; Indels 8; Gaps 1;  
 Qy 1 EVMLVESGGGLVQPGKSLKSCAASGFTFNMYAMNVRQAPGKLEWVARIRSKSNYYAT 60  
 Db 1 EVKLEESGGGLVQPGGSMKLSVASGFTFSNYMNMVRSQPEKGLWVAEIRLKS-HYAT 60  
 Qy 61 FYADSVKDRFTISRDDQSOMLYLQNMNLTEDTAMYYCVRGASGIDYAMDYWGQGTSLT 120  
 Db 61 HYAESVKGRTISRDDSKSVYLQNMNLRADDTGIYCTTGFA-----YWGQGTSLT 112  
 Qy 121 V 121  
 Db 113 V 113  
 RESULT 5  
 HV33\_MOUSE  
 ID HV33\_MOUSE STANDARD; PRT; 115 AA.  
 AC P01802;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region W3082.  
 OS Mus musculus (Mouse).  
 Query Match 67.8%; Score 438; DB 1; Length 113;  
 Best Local Similarity 70.2%; Pred. No. 4.5e-40;  
 Matches 85; Conservative 10; Mismatches 18; Indels 8; Gaps 1;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=82099361; PubMed=6798111;  
 RA Johnson N., Slankard J., Paul L., Hood L.;  
 RT "The complete V domain amino acid sequences of two myeloma inulin-  
 binding proteins";  
 RL J. Immunol. 128:302-307(1982).

CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT

CC BINDS INULIN.  
 DR PIR: D92811; AVMS82.  
 DR HSSP: P01810; 2FBJ.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 KW Immunoglobulin V region.  
 FT DISULFID 22 98  
 FT NON\_TER 115 115  
 SQ SEQUENCE 115 AA; 12887 MW; 9B4517648C121C5A CRC64;

Query Match 67.5%; Score 436; DB 1; Length 115;  
 Best Local Similarity 69.1%; Pred. No. 7.6e-40;  
 Matches 85; Conservative 12; Mismatches 18; Indels 8; Gaps 1;

Oy 1 EVMLVESGGGLVQPKGSLKSCAASGFTFNMYAMWVRQAPGKGLWVARIKSNVAT 60  
 Db 1 EVKLEESGGGLVQPGSKMLSCVAGSFTFSYMWVVRQSPKGLWVAEIRLKSNNVAT 60  
 Oy 61 FYADSVKDRFTISRDDSSMLYLQNNLKTEDTAMYYCVRRGASGIDYAMDYWGOGTSLT 120  
 Db 61 HYAESVKGRTISRDDSSVYLQNNLRPDTGIYCTTGA-----YWGOGTLVT 112

Oy 121 VSS 123

Db 113 VSA 115

RESULT 6

HV28\_MOUSE  
 ID HV28\_MOUSE STANDARD; PRT; 113 AA.  
 AC P01797;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region 061.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE.  
 RX MEDLINE=78158406; PubMed=417344;  
 RA Vrana M., Rudikoff S., Potter M.;  
 RT "Sequence variation among heavy chains from inulin-binding myeloma  
 proteins";  
 RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).

CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT

CC BINDS INULIN.  
 DR PIR: B93818; AVMS61.  
 DR HSSP: P01810; 2FBJ.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 KW Immunoglobulin V region.  
 FT DISULFID 22 98  
 FT NON\_TER 113 113  
 SQ SEQUENCE 113 AA; 12671 MW; 76634C121C598285 CRC64;

Query Match 66.6%; Score 430; DB 1; Length 113;

Best Local Similarity 69.4%; Pred. No. 3.3e-39;  
 Matches 84; Conservative 10; Mismatches 19; Indels 8; Gaps 1;

Oy 1 EVMLVESGGGLVQPKGSLKSCAASGFTFNMYAMWVRQAPGKGLWVARIKSNVAT 60  
 Db 1 EVKLEESGGGLVQPGSKMLSCVAGSFTFSYMWVVRQSPKGLWVAEIRLKSNNVAT 60  
 Oy 61 FYADSVKDRFTISRDDSSMLYLQNNLKTEDTAMYYCVRRGASGIDYAMDYWGOGTSLT 120  
 Db 61 HYAESVKGRTISRDDSSVYLQNNLRPDTGIYCTTGA-----YWGOGTLVP 112

Oy 121 V 121

Db 113 V 113

RESULT 7

HV01\_RAT  
 ID HV01\_RAT STANDARD; PRT; 142 AA.  
 AC P01805;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V region IR2 precursor.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=83064537; PubMed=6292865;  
 RA Hellman L., Pettersson U., Engstrom A., Karlsson T., Bennich H.;  
 RT "Structure and evolution of the heavy chain from rat immunoglobulin  
 E";  
 RL Nucleic Acids Res. 10:6041-6049(1982).  
 CC -1- MISCELLANEOUS: THE MRNA WAS ISOLATED FROM AN IGE-SECRETING  
 CC IMMUNOCYTOMA THAT ARISES SPONTANEOUSLY IN LOU/C/WSL RATS.

DR PIR: A02075; EVTR2.

DR HSSP: P01789; 1MCP.

DR InterPro: IPR003006; Ig\_MHC.

DR InterPro: IPR003596; Ig\_v.

DR Pfam: PF00047; Ig; 1.

DR SMART: SM00406; IGV; 1.

KW Immunoglobulin V region; Signal.

FT SIGNAL 1 19

FT CHAIN 20 142

FT NON\_TER 142 142

SQ SEQUENCE 142 AA; 16024 MW; DE29B6CFE745DF3B CRC64;

Query Match 66.1%; Score 427; DB 1; Length 142;

Best Local Similarity 65.9%; Pred. No. 8.9e-39;  
 Matches 83; Conservative 18; Mismatches 19; Indels 6; Gaps 3;

Oy 1 EVMLVESGGGLVQPKGSLKSCAASGFTFNMYAMWVRQAPGKGLWVARIKSNVAT 60  
 Db 20 EVKLEESGGGLVQPGMSVKLSATSGFTFSDYMWVRQAPGKGLWVAEIRKANNVVA 79

Oy 61 FYADSVKDRFTISRDDSSMLYLQNNLKTEDTAMYYCVRRGASGIDYAMDYWGOGT 117  
 Db 80 YGKSLKGRFTLSRDDSSVYLQNNLRPDTGIYCTTGA-----YSENFVYWGOGT 136

Oy 118 SLTVSS 123

Db 137 LVTVSS 142

RESULT 8

HV29\_MOUSE  
 ID HV29\_MOUSE STANDARD; PRT; 113 AA.  
 AC P01798;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig heavy chain V-III region E109.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=78158406; PubMed=417344;  
RA Vrana M., Rudikoff S., Potter M.;  
RT "Sequence variation among heavy chains from inulin-binding myeloma  
proteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).  
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
BINDS INULIN.  
DR PIR: C93818; AVMS09.  
DR HSSP: P01810; 2FBU.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; Igv; 1.  
KW Immunoglobulin V region.  
FT DISULFID 22 98  
FT NON\_TER 113 113 BY SIMILARITY.  
SQ SEQUENCE 113 AA: 12647 MW; EE50F2F20EDB129B CRC64;  
  
Query Match 65.9%; Score 426; DB 1; Length 113;  
Best Local Similarity 67.8%; Pred. No. 8.7e-39;  
Matches 82; Conservative 13; Mismatches 18; Indels 8; Gaps 1;  
  
QY 1 EYVLVESGGGLVQPGKSLKCAASGFTFNYYAMNVRQAPGKGLKLEWVARIRSKSNYYAT 60  
DB 1 EVKLEESGGGLVQPGGSKLSCVAGSFTFSNYMNMVVRQSPKGLKLEWVAEIRLKSNNYAT 60  
  
QY 61 FYADSVKDRFTISRDSQSMVLYOMNNLKTEDTANYCYVRRGASGIDYAMDYWGQTSLT 120  
DB 61 HYAESYKGRFTISRDSQSMVLYOMNNLKRAEDTGIHYCTGFA-----YWGQGTLYT 112  
  
QY 121 V 121  
DB 113 V 113  
  
RESULT 9  
HV31\_MOUSE  
ID HV31\_MOUSE STANDARD; PRT; 113 AA.  
AC P01800;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region T957.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=81216632; PubMed=6787122;  
RA Rudikoff S., Potter M.;  
RT "Immunoglobulin heavy chains from anti-inulin myeloma proteins:  
evidence for a new heavy chain joining segment.";  
RL J. Immunol. 127:191-194(1981).  
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
BINDS INULIN.  
DR PIR: A92810; AVMS57.  
DR HSSP: P01810; 2FBU.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; Igv; 1.  
KW Immunoglobulin V region.  
FT DISULFID 22 98  
FT NON\_TER 113 113 BY SIMILARITY.  
SQ SEQUENCE 113 AA: 12732 MW; 26618F626B59859E CRC64;

Ig heavy chain V-III region E109.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=78158406; PubMed=417344;  
RA Vrana M., Rudikoff S., Potter M.;  
RT "Sequence variation among heavy chains from inulin-binding myeloma  
proteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).  
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
BINDS INULIN.  
DR PIR: C93818; AVMS09.  
DR HSSP: P01810; 2FBU.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; Igv; 1.  
KW Immunoglobulin V region.  
FT DISULFID 22 98  
FT NON\_TER 113 113 BY SIMILARITY.  
SQ SEQUENCE 113 AA: 12647 MW; EE50F2F20EDB129B CRC64;  
  
Query Match 65.9%; Score 426; DB 1; Length 113;  
Best Local Similarity 67.8%; Pred. No. 8.7e-39;  
Matches 82; Conservative 13; Mismatches 18; Indels 8; Gaps 1;  
  
QY 1 EYVLVESGGGLVQPGKSLKCAASGFTFNYYAMNVRQAPGKGLKLEWVARIRSKSNYYAT 60  
DB 1 EVKLEESGGGLVQPGGSKLSCVAGSFTFSNYMNMVVRQSPKGLKLEWVAEIRLKSNNYAT 60  
  
QY 61 FYADSVKDRFTISRDSQSMVLYOMNNLKTEDTANYCYVRRGASGIDYAMDYWGQTSLT 120  
DB 61 HYAESYKGRFTISRDSQSMVLYOMNNLKRAEDTGIHYCTGFA-----YWGQGTLYT 112  
  
QY 121 V 121  
DB 113 V 113  
  
RESULT 9  
HV31\_MOUSE  
ID HV31\_MOUSE STANDARD; PRT; 113 AA.  
AC P01800;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region T957.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=81216632; PubMed=6787122;  
RA Rudikoff S., Potter M.;  
RT "Immunoglobulin heavy chains from anti-inulin myeloma proteins:  
evidence for a new heavy chain joining segment.";  
RL J. Immunol. 127:191-194(1981).  
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
BINDS INULIN.  
DR PIR: A92810; AVMS57.  
DR HSSP: P01810; 2FBU.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; Igv; 1.  
KW Immunoglobulin V region.  
FT DISULFID 22 98  
FT NON\_TER 113 113 BY SIMILARITY.  
SQ SEQUENCE 113 AA: 12732 MW; 26618F626B59859E CRC64;

Query Match 65.3%; Score 422; DB 1; Length 113;  
Best Local Similarity 68.6%; Pred. No. 2.3e-38;  
Matches 83; Conservative 10; Mismatches 20; Indels 8; Gaps 1;  
  
QY 1 EYVLVESGGGLVQPGKSLKCAASGFTFNYYAMNVRQAPGKGLKLEWVARIRSKSNYYAT 60  
DB 1 EVKLEESGGGLVQPGGSKLSCVAGSFTFSNYMNMVVRQSPKGLKLEWVAEIRLKSNNYAT 60  
  
QY 61 FYADSVKDRFTISRDSQSMVLYOMNNLKTEDTANYCYVRRGASGIDYAMDYWGQTSLT 120  
DB 61 HYAESYKGRFTISRDSQSMVLYOMNNLKRAEDTGIHYCTGFA-----YWGQGTLYT 112  
  
QY 121 V 121  
DB 113 V 113  
  
RESULT 10  
HV25\_MOUSE  
ID HV25\_MOUSE STANDARD; PRT; 123 AA.  
AC P01794;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V region HPGC14.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=81197602; PubMed=7231520;  
RA Gearhart P.J., Johnson N.D., Douglas R., Hood L.;  
RT "IgG antibodies to phosphorylcholine exhibit more diversity than  
their IgM counterparts.";  
RL Nature 291:29-34(1981).  
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
BINDS PHOSPHORYLCHOLINE.  
DR PIR: A02070; AVMS75.  
DR HSSP: P01789; 1MCP.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; Igv; 1.  
KW Immunoglobulin V region; Hybridoma.  
FT NON\_TER 123 123  
SQ SEQUENCE 123 AA: 13807 MW; A7584FB098B7785D CRC64;  
  
Query Match 65.2%; Score 421; DB 1; Length 123;  
Best Local Similarity 65.9%; Pred. No. 3.3e-38;  
Matches 81; Conservative 14; Mismatches 28; Indels 0; Gaps 0;  
  
QY 1 EYVLVESGGGLVQPGKSLKCAASGFTFNYYAMNVRQAPGKGLKLEWVARIRSKSNYYAT 60  
DB 1 EVKLEESGGGLVQPGGSKLSCATSGTFSAFYMEVVRQPPGKRLKLEIAASRKANDYTT 60  
  
QY 61 FYADSVKDRFTISRDSQSMVLYOMNNLKTEDTANYCYVRRGASGIDYAMDYWGQTSLT 120  
DB 61 EYSASVYKGRFFVSRTDSQSLYLMNNALRAEDTAIIYCARDYVGYDFWVGAGTTT 120  
  
QY 121 VSS 123  
DB 121 VSS 123  
  
RESULT 11  
HV16\_MOUSE  
ID HV16\_MOUSE STANDARD; PRT; 136 AA.  
AC P01783;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)









```
RESULT 2
Q8WU38
ID Q8WU38 PRELIMINARY; PRT; 573 AA.
AC Q8WU38;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 63.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021276; AAH21276.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00409; IG; 1.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 573 AA; 62967 MW; FD072344033AC530 CRC64;
Query Match 66.3%; Score 428.5; DB 4; Length 573;
Best Local Similarity 66.7%; Pred. No. 3.7e-36;
Matches 84; Conservative 17; Mismatches 20; Indels 5; Gaps 2;

QY 1 EYVLVESGGGLVQPGKSLKLSAASGFTFNNTYAMNVRQAPGKLEWVARIRSKSNYYAT 60
  || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 20 EVLVESGGGLVQPGKSLKLSAASGFTDDYAMHVRQAPGKLEWVGSIWSSGISG- 78

QY 61 FYADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYCYVRRGASGIDYAMDYWGQTSLT 117
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 79 -YADSVKGRFTISRDNKNSLYLQNNLSRAEDTALYYCAKHGSGSYIGYYGMDVWGQGT 137

QY 118 STVSS 123
  ::|||
DB 138 TTVSS 143

RESULT 3
Q9UL90
ID Q9UL90 PRELIMINARY; PRT; 113 AA.
AC Q9UL90;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
  fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035023; AAD56259.1; -
DR HSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;
Query Match 64.9%; Score 419.5; DB 4; Length 118;
Best Local Similarity 68.0%; Pred. No. 4.4e-36;
Matches 83; Conservative 15; Mismatches 19; Indels 5; Gaps 2;

QY 1 EYVLVESGGGLVQPGKSLKLSAASGFTFNNTYAMNVRQAPGKLEWVARIRSKSNYYAT 60
  || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 EVLVESGGGLVQPGKSLKLSAASGFTFSYSMNVRQAPGKLEWVSYI--SSTIIFI 58

QY 61 FYADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYCYVRRGASGIDYAMDYWGQTSLT 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 59 YIADSVKGRFTISRDNKNSLYLQNNLSRAEDTAVYYCAK-----DLNYWGQGTLT 110

QY 121 VSS 123
  |||
DB 111 VSS 113

RESULT 4
Q9UL91
ID Q9UL91 PRELIMINARY; PRT; 118 AA.
AC Q9UL91;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
  fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035023; AAD56259.1; -
DR HSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;
Query Match 64.9%; Score 419.5; DB 4; Length 118;
Best Local Similarity 68.0%; Pred. No. 4.4e-36;
Matches 83; Conservative 15; Mismatches 19; Indels 5; Gaps 2;

QY 1 EYVLVESGGGLVQPGKSLKLSAASGFTFNNTYAMNVRQAPGKLEWVARIRSKSNYYAT 60
  || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 EVLVESGGGLVQPGKSLKLSAASGFTFSYSMNVRQAPGKLEWVSYI--SSTIIFI 58

QY 61 FYADSVKDRFTISRDSQSMLYLQNNLKTEDTAMYCYVRRGASGIDYAMDYWGQTSLT 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 59 YIADSVKGRFTISRDNKNSLYLQNNLSRAEDTAVYYCAK-----EAFDYGQGTWT 115

QY 121 VS 122
  ||
DB 116 VS 117

RESULT 5
```



```
RP SEQUENCE FROM N.A.
RC TISSUE-MAMMARY GLAND;
RA Isozai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Takatsuma M., Hosofiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027379; BAB5072.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 494 AA; 5308 MW; 9A1D7AEBSAEE4C0E CRC64;

Query Match 64.0%; Score 413.5; DB 4; Length 494;
Best Local Similarity 65.3%; Pred. No. 1.1e-34;
Matches 81; Conservative 18; Mismatches 22; Indels 3; Gaps 2;

QY 1 EYMLVSGGLVQPKGSLKSLCAASGFTFNYYAMNVRQAPGKGLWVARIKSKSNYYAT 60
DB 20 EVOLVESGGGLVQPKGSLKSLCAASGLSFSTYAMNVRQAPGKGLWVARIKSKSNYYAT 77
QY 61 FYADSVKDRFTISRDDSQSLYLMNNLTKEDTAMYYCVRRGASG-IDYAMDYWGQGTSL 119
DB 78 YYRDSVKGRFTISRDNKNSLYLMNLSLRVDDTAVYYCARDSCNGAICYGFSPWGQGTIV 137
QY 120 TVSS 123
DB 138 TVSS 141

RESULT 9
Q8WUK1 PRELIMINARY; PRT; 613 AA.
AC Q8WUK1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 67.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FOHSIL;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020240; AAH20240.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;

Query Match 64.0%; Score 413.5; DB 4; Length 613;
Best Local Similarity 65.0%; Pred. No. 1.5e-34;
Matches 80; Conservative 19; Mismatches 21; Indels 3; Gaps 2;

QY 1 EYMLVSGGLVQPKGSLKSLCAASGFTFNYYAMNVRQAPGKGLWVARIKSKSNYYAT 60
DB 20 QVOLVESGGVQPKGRSLKSLCAASGFTFSSYGMHVRQAPGKGLWVAVISYDGSN--K 77
QY 61 FYADSVKDRFTISRDDSQSLYLMNNLTKEDTAMYYCVRRGASGIDYAMDYWGQGTSLT 120
```

```
DB 78 YYADSVKGRFTISRDNKNSLYLMNLSLRVDDTAVYYCARDSCNGAICYGFSPWGQGTIVT 136
QY 121 VSS 123
DB 137 VSS 139

RESULT 10
Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update).
DE Single chain Fv (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RT "An antibody fragment2A3 specific for native lysozyme : isolation from a
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049915; BAB16829.1; -.
DR HSP; P01772; 2FB4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match 63.7%; Score 411.5; DB 4; Length 112;
Best Local Similarity 67.5%; Pred. No. 2.8e-35;
Matches 79; Conservative 17; Mismatches 14; Indels 7; Gaps 2;

QY 1 EYMLVSGGLVQPKGSLKSLCAASGFTFNYYAMNVRQAPGKGLWVARIKSKSNYYAT 60
DB 1 EVOLVESGGGVVQPKGSLKSLCAASGFTFDDYGMHVRQAPGKGLWVSGI--NNNGGST 58
QY 61 FYADSVKDRFTISRDDSQSLYLMNNLTKEDTAMYYCVRRGASGIDYAMDYWGQGT 117
DB 59 GYADSVKGRFTISRDNKNSLYLMNLSLRVDDTAVYYCARR-----RYALDYWGQGT 110

RESULT 11
Q8TC77 PRELIMINARY; PRT; 471 AA.
AC Q8TC77;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 51.8 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024289; AAH24289.1; -.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;
```



```
Db 1 EVLVESGGVQVPGRSLRSLCAASRFTFSNYGMHWVRQAPGKLEWVAASNDGSN--K 58
QY 61 FYADSVKDRFTISRDDSQSMYLIQMNHLKTEDTAMYYCVR--RGASSIDYAMDYWGQGT 118
Db 59 FYADSVKGRFTIFRDNKSNMMDLQMSLRRAEDTAVYICARDERGLVGY-FDYWGQGT 117
QY 119 LTVSS 123
Db 118 VTVSS 122
```

RESULT 15

```
Q91WPS
ID Q91WP5 PRELIMINARY: PRT: 479 AA.
AC Q91WP5:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical 51.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013656; AAH13656.1;
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 479 AA; 51603 MW; ECB2D087748584F CRC64;
```

```
Query Match 62.4%; Score 403; DB 11; Length 479;
Best Local Similarity 64.2%; Pred. No. 1.3e-33;
Matches 79; Conservative 18; Mismatches 18; Indels 8; Gaps 2;

QY 1 EVMLVESGGGLVQPKGSLKSLCAASGFTFNMYAMNVRQAPGKLEWVARIRKSNYYAT 60
Db 20 EVLVESGGGLVQPKGSLKSLCAASGLTFSNYAMSVMWVRQSPKRLWVAASNSNGN--T 77
QY 61 FYADSVKDRFTISRDDSQSMYLIQMNHLKTEDTAMYYCVRRGASGIDYAMDYWGQGTSLT 120
Db 78 YSDTKGRFTISRDNKSTLYLQMSLSRSEDYFYCVRRG-----YFDVWGAGTAVT 131
QY 121 VSS 123
Db 132 VSS 134
```

Search completed: November 27, 2002, 07:22:53  
Job time : 74.3861 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 27, 2002, 07:17:51 : Search time 22.3205 Seconds  
(without alignments)  
162.139 Million cell updates/sec

Title: US-09-893-615-87

Perfect score: 646

Sequence: 1 EVMLVSGGLVQPKGSLK.....SGIDYMDYMGQTSLSVSS 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCRUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555	85.9	115	3	US-08-767-128-36
2	549	85.0	119	1	US-08-442-542-2
3	549	85.0	119	3	US-08-765-469-2
4	549	85.0	599	1	US-08-442-542-18
5	549	85.0	599	3	US-08-765-469-18
6	532	82.4	110	3	US-08-767-128-24
7	462.5	71.6	122	3	US-08-483-749A-2
8	460.5	71.3	120	4	US-09-025-769B-38
9	460.5	71.3	120	4	US-09-025-769B-63
10	460.5	71.3	281	4	US-09-025-769B-178
11	454.5	70.4	120	3	US-08-767-128-26
12	446	69.0	119	3	US-08-767-128-26
13	444.5	68.8	114	3	US-08-483-749A-10
14	444.5	68.8	124	3	US-08-767-128-38
15	444.5	68.8	263	4	US-09-069-821-3
16	444.5	68.8	283	4	US-09-420-592A-6
17	444	68.7	117	4	US-09-025-769B-24
18	443.5	68.7	130	1	US-08-478-039-70
19	443.5	68.7	130	1	US-08-478-039-93
20	443.5	68.7	130	1	US-08-476-349A-70
21	443.5	68.7	130	1	US-08-476-349A-93
22	441.5	68.3	120	2	US-07-934-373C-4
23	441.5	68.3	120	3	US-08-437-642B-4
24	441.5	68.3	120	4	US-08-146-206C-4
25	441.5	68.3	140	4	US-08-983-607-32
26	441.5	68.3	250	2	US-08-392-338A-15
27	441.5	68.3	250	3	US-09-166-750-15

Sequence 15, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 14, Appl  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 13, Appl  
Sequence 13, Appl

## ALIGNMENTS

### RESULT 1

US-08-767-128-36  
Sequence 36, Application US/08767128  
Patent No. 6111079  
GENERAL INFORMATION:  
APPLICANT: WYLIE, DWANE E.  
APPLICANT: LOPEZ, OSVALDO  
APPLICANT: MURRAY, PETER JOSEPH  
APPLICANT: GOEBEL, PETER  
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 6111079west Center, 90 South Seventh St  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,128  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE: 04-DEC-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09258  
FILING DATE: 05-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/541,373  
FILING DATE: 10-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/462,798  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Carter, Charles G.  
REGISTRATION NUMBER: 35,093  
REFERENCE/DOCKET NUMBER: 8648.49USF1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612/371-5278  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:

US-08-767-128-36

Query Match 85.9%; Score 555; DB 3; Length 115;  
Best Local Similarity 87.8%; Pred. No. 1.6e-49;  
Matches 108; Conservative 2; Mismatches 5; Indels 8; Gaps 1;  
  
QY 1 EYVLVESGGGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPKGLWVARIRSKSNYYAT 60  
|||  
DB 1 EYVLVESGGGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPKGLWVARIRSKSNYYAT 60  
  
QY 61 FYADSVKDRFTISRDDSQSMFLQMNLLKTEDTAMYCYVRRGASGIDYAMDYWGQGTSLT 120  
|||  
DB 61 FYADSVKDRFTISRDDSQSMFLQMNLLKTEDTAMYCYVRRGASGIDYAMDYWGQGTSLT 112  
  
QY 121 VSS 123  
|||  
DB 113 VSS 115

RESULT 2  
US-08-442-542-2  
; Sequence 2, Application US/08442542  
; Patent No. 5686600  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine B.  
; APPLICANT: Koziel, Michael G.  
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
; TITLE OF INVENTION: Proteins and their Use  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ciba-Geigy Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,542  
FILING DATE: 16-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/267,641  
FILING DATE: 28-JUN-1994  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-442-542-2

Query Match 85.0%; Score 549; DB 1; Length 119;  
Best Local Similarity 86.3%; Pred. No. 6.8e-49;  
Matches 107; Conservative 5; Mismatches 6; Indels 6; Gaps 2;

QY 1 EYVLVESGGGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPKGLWVARIRSKSNYYAT 60  
:|||  
DB 1 QVKLQESGGGLVQPKGSLKLSCAASGFTFNFMNVRQAPKGLWVARIRSKSNYYAT 60  
  
QY 61 FYADSVKDRFTISRDDSQSMFLQMNLLKTEDTAMYCYVRRGASGIDYAMDYWGQGTSL 119  
|||  
DB 61 SYGDSVKDRFTVSRDDSSQSMFLQMNLLKTEDTAMYCYVRRGASGIDYAMDYWGQGTSLV 115  
  
QY 120 TVSS 123  
|||  
DB 116 TVSS 119

RESULT 3  
US-08-765-469-2  
; Sequence 2, Application US/08765469  
; Patent No. 6069301  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine B.  
; APPLICANT: Koziel, Michael G.  
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
; TITLE OF INVENTION: Proteins and their Use  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ciba-Geigy Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/765,469  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/267,641  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-765-469-2

Query Match 85.0%; Score 549; DB 3; Length 119;  
Best Local Similarity 86.3%; Pred. No. 6.8e-49;  
Matches 107; Conservative 5; Mismatches 6; Indels 6; Gaps 2;

QY 1 EYVLVESGGGLVQPKGSLKLSCAASGFTFNMYAMNVRQAPKGLWVARIRSKSNYYAT 60  
:|||  
DB 1 QVKLQESGGGLVQPKGSLKLSCAASGFTFNFMNVRQAPKGLWVARIRSKSNYYAT 60  
  
QY 61 FYADSVKDRFTISRDDSQSMFLQMNLLKTEDTAMYCYVRRGASGIDYAMDYWGQGTSL 119  
|||  
DB 61 SYGDSVKDRFTVSRDDSSQSMFLQMNLLKTEDTAMYCYVRRGASGIDYAMDYWGQGTSLV 115



COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,128  
FILING DATE: 05-JUN-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09258  
FILING DATE: 05-JUN-1996  
APPLICATION NUMBER: 08/541,373  
FILING DATE: 10-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/462,798  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Carter, Charles G.  
REGISTRATION NUMBER: 35,093  
REFERENCE/DOCKET NUMBER: 8648.49USF1  
TELEPHONE: 612/371-5278  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
US-08-767-128-24

Query Match 82.4%; Score 532; DB 3; Length 110;  
Best Local Similarity 87.3%; Pred. No. 3.3e-47;  
Matches 103; Conservative 3; Mismatches 4; Indels 8; Gaps 1;  
QY 6 EGGGLVQPKGSLKLSAASGFTFNMYAMVVRQAPGKGLWVARIRSKSNYYATFYADS 65  
DB 1 EGGGLVQPKGSLKLSAASGFTFNMYAMVVRQAPGKGLWVARIRSKSNYYATFYADS 60  
QY 66 VKDRFTISRDDSQSMYLYQNNLNKLTEDTAMYCVRRGASGIDYAMDYWGQGTSLTVSS 123  
DB 61 VKDRFTISRDDSQSMYLYQNNLNKLTEDTAMYCVRRFA-----YWAQGTILVTYSA 110  
RESULT 7  
US-08-483-749A-2  
Sequence 2, Application US/08483749A  
Patent No. 6034561  
GENERAL INFORMATION:  
APPLICANT: RING, DAVID B.  
TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY  
TITLE OF INVENTION: MOLECULES SPECIFIC FOR CANCER ANTIGENS  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHIRON CORPORATION  
STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097  
CITY: EMERYVILLE  
STATE: CA  
COUNTRY: USA  
ZIP: 94662-8097  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,749A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: SAVEREIDE, PAUL B.  
REGISTRATION NUMBER: 36,914  
REFERENCE/DOCKET NUMBER: 0508.008  
TELEPHONE: (510) 601-2585  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-483-749A-2  
Query Match 71.6%; Score 462.5; DB 3; Length 122;  
Best Local Similarity 73.2%; Pred. No. 4.4e-40;  
Matches 90; Conservative 10; Mismatches 22; Indels 1; Gaps 1;  
QY 1 EYMLVESGGGLVQPKGSLKLSAASGFTFNMYAMVVRQAPGKGLWVARIRSKSNYYAT 60  
DB 1 EVKLEESGGGLVQPKGSRMKLSVASGFTFSYWMNVVRQAPGKGLWVARIRSKSNYYAT 60  
QY 61 FYADSVKDRFTISRDDSQSMYLYQNNLNKLTEDTAMYCVRRGASGIDYAMDYWGQGTSLT 120  
DB 61 HYAESVKGREFTISRDDSQSMYLYQNNLNKLTEDTAMYCVRRGASGIDYAMDYWGQGTSLT 119  
QY 121 VSS 123  
DB 120 VSS 122  
RESULT 8  
US-09-025-769B-38  
Sequence 38, Application US/09025769B  
Patent No. 6300064  
GENERAL INFORMATION:  
APPLICANT: Knappik, Achim  
APPLICANT: Pack, Peter  
APPLICANT: Ilag, Vic  
APPLICANT: Ge, Liming  
APPLICANT: Moroney, Simon  
APPLICANT: Plueckthun, Andreas  
TITLE OF INVENTION: Protein/(Poly)peptide libraries  
NUMBER OF SEQUENCES: 373  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10021  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/025,769B  
FILING DATE: 18-FEB-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95 11 3021.0  
FILING DATE: 18-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: James F. Haley, Jr., Esq.

```
;
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-38

Query Match 71.3%; Score 460.5; DB 4; Length 120;
Best Local Similarity 72.4%; Pred. No. 6.9e-40;
Matches 89; Conservative 14; Mismatches 17; Indels 3; Gaps 2;

QY 1 EVMVLSGGGLVQPGKSLKLSCAASGFTFNMYAMNVRQAPGKGLWVARIRKSNYYAT 60
   ||||||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 EVOLVESGGGLVQPGKSLKLSCAASGFTFSSYAMNVRQAPGKGLWVSATSGSGG--ST 58

QY 61 FYADSVKDRFTISRDSOSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYWGQGTSLT 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 YYADSVKGRFTISRDNKNTLYLQNNLSRAEDTAVYYCARWGGDGF-YAMDYWGQGTTLVT 117

QY 121 VSS 123
   |||
Db 118 VSS 120

RESULT 9
US-09-025-769B-63
; Sequence 63, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025.769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
US-09-025-769B-178

Query Match 71.3%; Score 460.5; DB 4; Length 281;
Best Local Similarity 72.4%; Pred. No. 1.8e-39;
Matches 89; Conservative 14; Mismatches 17; Indels 3; Gaps 2;

QY 1 EVMVLSGGGLVQPGKSLKLSCAASGFTFNMYAMNVRQAPGKGLWVARIRKSNYYAT 60
   ||||||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
;
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-63

Query Match 71.3%; Score 460.5; DB 4; Length 120;
Best Local Similarity 72.4%; Pred. No. 6.9e-40;
Matches 89; Conservative 14; Mismatches 17; Indels 3; Gaps 2;

QY 1 EVMVLSGGGLVQPGKSLKLSCAASGFTFNMYAMNVRQAPGKGLWVARIRKSNYYAT 60
   ||||||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 EVOLVESGGGLVQPGKSLKLSCAASGFTFSSYAMNVRQAPGKGLWVSATSGSGG--ST 58

QY 61 FYADSVKDRFTISRDSOSMLYLQNNLKTEDTAMYYCVRRCASGIDYAMDYWGQGTSLT 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 YYADSVKGRFTISRDNKNTLYLQNNLSRAEDTAVYYCARWGGDGF-YAMDYWGQGTTLVT 117

QY 121 VSS 123
   |||
Db 118 VSS 120

RESULT 10
US-09-025-769B-178
; Sequence 178, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025.769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-178
```

```
Db 26 EVLVESGGGLVQPGKSLRSLSCAASGFTFSYAMSWVRQAPGKLEWVSAISGGG--ST 83
QY 61 FYADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVRRGASGIDYAMDYWGQGTSLT 120
Db 84 YYADSVKGRFTISRDSKNTLYLQMSLRADTAVYICARWGGDGP-YANDYWGQGTSLT 142
QY 121 VSS 123
Db 143 VSS 145

RESULT 11
US-08-767-128-28
; Sequence 28, Application US/08767128
; Patent No. 6111079
; GENERAL INFORMATION:
; APPLICANT: WYLIE, DWANE E.
; APPLICANT: LOPEZ, OSVALDO
; APPLICANT: MURRAY, PETER JOSEPH
; APPLICANT: GOEBEL, PETER
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 NO. 6111079west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,128
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09258
; FILING DATE: 05-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/541,373
; FILING DATE: 10-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,798
; FILING DATE: 05-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Carter, Charles G.
; REGISTRATION NUMBER: 35,093
; REFERENCE/DOCKET NUMBER: 8648.49USF1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/371-5278
; TELEFAX: 612/332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; US-08-767-128-28
```

```
Query Match 70.4%; Score 454.5; DB 3; Length 120;
Best Local Similarity 70.7%; Pred. No. 2.8e-39;
Matches 87; Conservative 11; Mismatches 22; Indels 3; Caps 1;

QY 1 EVMVLVSGGGLVQPGKSLKSLSCAASGFTFNYYAMNVRQAPGKLEWVAIRSKSNYYAT 60
Db 1 EVKTEESGGGLVQPGGSKLSVASGFTFSNWNWVRQSPKGLWVAEIRLKSNNYAT 60
QY 61 FYADSVKDRFTISRDSQSMLYLQMNMLKTEDTAMYYCVRRGASGIDYAMDYWGQGTSLT 120
Db 61 HYAESVKGRTISRDSKSSVYLQMNMLRAEDTGIYYCTRYGRG---GAYWGEGLT 117
QY 121 VSS 123
Db 118 VSA 120

RESULT 12
US-08-767-128-26
; Sequence 26, Application US/08767128
; Patent No. 6111079
; GENERAL INFORMATION:
; APPLICANT: WYLIE, DWANE E.
; APPLICANT: LOPEZ, OSVALDO
; APPLICANT: MURRAY, PETER JOSEPH
; APPLICANT: GOEBEL, PETER
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 NO. 6111079west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,128
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09258
; FILING DATE: 05-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/541,373
; FILING DATE: 10-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,798
; FILING DATE: 05-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Carter, Charles G.
; REGISTRATION NUMBER: 35,093
; REFERENCE/DOCKET NUMBER: 8648.49USF1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/371-5278
; TELEFAX: 612/332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
QY      61  FYADSVKDRFTISRDDSQSMLYLYQMNLKTEDTAMYTCVRRGASGIDYANDYWGQTSLT 120
      ||:|'|' || |||||: :|||||:|||| :|| |||||:|:|
```

Matches 87; Conservative 12; Mismatches 24; Indels 1

QY 1 EVMLVESGGGLVQPKGSLKLSAASGTFNNYAMNWRQAPGKLEWVARIRSKSNYYAT 60  
Db 1 EVKLVESGGGLVQPGGSLRLSCLATSGTFTDYMSWVRQPGKALEWLGLIRNKANGYTT 60  
QY 61 FYADSVKDRFTTISRDDSQSMLYLQMNLLKTEDTAMYYCVRRGASGID-YAMDYWGQGTSL 119  
Db 61 EYSASVKGRTTISRDNQSILYLQMNLLRAEDSATYYCARDIYYDYDYAMDYWGQGTSV 120  
QY 120 TVSS 123  
Db 121 TVSS 124  
RESULT 15  
US-09-069-821-3  
; Sequence 3, Application US/09069821  
; Patent No. 6323322  
; GENERAL INFORMATION:  
; APPLICANT: FILPULA, DAVID  
; APPLICANT: WANG, MAOLIANG  
; APPLICANT: SHORR, ROBERT  
; APPLICANT: WHITLOW, MARC  
; APPLICANT: LEE, LIHSYNG S.  
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS  
; TITLE OF INVENTION: CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/069,821  
; FILING DATE: 30-APR-1998  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/067,341  
; FILING DATE: 02-DEC-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/063,074  
; FILING DATE: 27-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/050,472  
; FILING DATE: 23-JUN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/044,449  
; FILING DATE: 30-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KIM, JUDITH U.  
; REGISTRATION NUMBER: 40,679  
; REFERENCE/DOCKET NUMBER: 0977.2280003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)371-2600  
; TELEFAX: (202)371-2540  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 263 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-09-069-821-3

Query Match 68.8%; Score 444.5; DB 4; Length 263;  
Best Local Similarity 66.4%; Pred. No. 7.2e-38;  
Matches 87; Conservative 17; Mismatches 18; Indels 9; Gaps 2;

QY 1 EVMLVESGGGLVQPKGSLKLSAASGTFNNYAMNWRQAPGKLEWVARIRSKSNYYAT 60  
Db 130 EVQLVESGGGLVQPGGSLRLSCLATSGTFTDYMSWVRQPGKALEWLGLIRNKANGYTT 189  
QY 61 FYADSVKDRFTTISRDDSQSMLYLQMNLLKTEDTAMYYCVRRGASGID-----YAMDY 112  
Db 190 YYADSVKGRFTTISRDNKNTLYLQMNLLRAEDTAVYYCA-RGRXGXSLSGXYYYHYFDY 248  
QY 113 WGQGTSLTVSS 123  
Db 249 WGQGTSLTVSS 259

Search completed: November 27, 2002, 07:25:12  
Job time : 24.3205 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:36:38 ; Search time 1460.96 Seconds  
(without alignments)  
7350.598 Million cell updates/sec

Title: US-09-893-615-86  
Perfect score: 369  
Sequence: 1 GAAGTGATGCTGGTGGAGTC.....CCTCACTACCGTCCTCTCA 369

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl : \*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_or.\*
- 21: em\_ov.\*
- 22: em\_pat.\*
- 23: em\_ph.\*
- 24: em\_pl.\*
- 25: em\_ro.\*
- 26: em\_sts.\*
- 27: em\_un.\*
- 28: em\_vl.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sv.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	309.8	84.0	370	10	S65980	S65980 Ig VH193020
2	305	82.7	366	10	MUSIGH339A	L30140 Mus Musculu
3	304	82.4	475	10	MUSIGHAAAC	M60022 Mouse Ig he
4	303.4	82.2	499	10	MUSIGHAAAB	M60021 Mouse Ig he
5	301	81.6	443	10	MUSIGHXS	M21469 Mouse Ig un
6	301	81.6	444	10	MMIGHRF2	X16801 Mouse MRL-R
7	300.4	81.4	357	6	I73737	I73737 Sequence 1
8	300.4	81.4	1797	6	I73745	I73745 Sequence 17
9	298.6	80.9	366	10	AF178591	AF178591 Mus muscu
10	297	80.5	366	10	AF178595	AF178595 Mus muscu
11	297	80.5	366	10	AF178596	AF178596 Mus muscu
12	291	78.9	408	10	AF045483	AF045483 Mus muscu
13	289.6	78.5	366	10	MUSIGH125A	L31512 Mus musculu
14	288.8	78.3	360	10	MUSIGHCVRA	D50376 Mouse mRNA
15	288.6	78.2	345	6	ARI08686	ARI08686 Sequence
16	287.8	78.0	351	10	AF064445	AF064445 Mus muscu
17	286.2	77.6	663	10	AF064445	AF064445 Mus muscu
18	285.8	77.5	474	10	MUSIGHKCLT	M20829 Mouse Ige2a
19	284.4	77.1	330	6	ARI08680	ARI08680 Sequence
20	283.8	76.9	375	6	AX025327	AX025327 Sequence
21	283.4	76.8	185591	2	AC073561	AC073561 Mus muscu
22	283.4	76.8	185591	2	AC073561	AC073561 Mus muscu
23	282.4	76.5	661	10	AF064446	AF064446 Mus muscu
24	282.2	76.5	357	10	MUSIGHUHF	M5948 Mouse Ig ac
25	280.6	76.0	348	10	AF468835	AF468835 Mus muscu
26	279.4	75.7	360	10	AF006574	AF006574 Mus muscu
27	279.2	75.7	357	10	MUSIGHXT	M21470 Mouse Ig ac
28	279.2	75.7	360	10	MDIGMVAQ	222080 M.domesticu
29	279	75.6	316	10	AF118969	AF118969 Mus muscu
30	277.8	75.3	375	10	S63182	S63182 Ig VH-anti-
31	276.4	74.9	353	10	AF321939	AF321939 Mus muscu
32	276.2	74.9	421	10	MMIGHIHV	X78107 M.musculus
33	276	74.8	360	10	MDIGMVAR	222082 M.domesticu
34	274.4	74.4	202704	10	AC090843	AC090843 Mus muscu
35	274	74.3	312	10	MUSIGHACT	M34735 Mouse Ig H-
36	273.8	74.2	337	10	MMU55546	U55546 Mus musculu
37	273.2	74.0	337	10	MMU55545	U55545 Mus musculu
38	272.8	73.9	663	10	AF064442	AF064442 Mus muscu
39	272.8	73.9	663	10	AF064444	AF064444 Mus muscu
40	272.4	73.8	783	10	MMSCFVCP	Z70661 Artificial
41	269.6	73.1	360	10	MDIGGVAT	222093 M.domesticu
42	269.2	73.0	443	6	AX208064	AX208064 Sequence
43	269.2	73.0	443	6	AX208067	AX208067 Sequence
44	264.8	71.8	381	10	S72511	S72511 anti-estradi
45	264.6	71.7	333	10	MMU55544	U55544 Mus musculu

ALIGNMENTS

RESULT 1  
S65980  
LOCUS  
DEFINITION  
Ig VH193020-anti-insulin 193020 monoclonal antibody heavy chain  
variable region [mice, NOD spleen and myeloma cells, nonobese  
diabetic, mRNA Partial, 370 nt].  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
AUTHORS  
1 (bases 1 to 370)  
Pleau,J.M., Marche,P.N., Serrano,M.P., Boitard,C. and Bach,J.F.

SOURCE	ORGANISM	CDNA TO mRNA
Mouse (strain C57BL/6)	Mouse	spleen cell hybridoma 222
	Mus musculus	

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Brigido.M.M. and Stollar,B.D.
Two induced anti-z-DNA monoclonal antibodies use VH gene segments
related to those of anti-DNA autoantibodies
J. Immunol. 146 (6), 2005-2009 (1991)
91170743
PUBMED
1900879
FEATURES
    source
        Location/Qualifiers
            1..475
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /map="chromosome 12"
                /cell_line="hybridoma 222"
                /tissue_type="liver"
                /dev_stage="fetus"
            1..475
                /gene="IgH"
            53..475
                /partial
            /gene="IgH"
            /codon_start=1
            /product="immunoglobulin heavy chain"
            /db_xref="taxon:10090"
            /protein_id="AAAS1056.1"
            /db_xref="GI:194551"
            /translation="MLGLKVVFFVYQGVHCKVQLVESGGGLVQPKGSLKLSCAAS
            GFNFNTYAMNVROAPGKLEWVARIRKSNVATYADSKDRFTISRDDSENMLYL
            QNINLKAEDTAMYICVRQAYSNGAMDYWGQISVTSS"
            53..109
                /gene="IgH"
            110..475
                /partial
            /gene="IgH"
            /product="immunoglobulin heavy chain"
            110..409
                /gene="IgH"
            410..424
                /gene="IgH"
            425..475
                /gene="IgH"
            /gene="IgH"
BASE COUNT    131 a    91 c    114 g    139 t
ORIGIN          Chromosome 12.
Query Match    82.4%; Score 304; DB 10; Length 475;
Best Local Similarity 91.0%; Pred. No. 5.5e-84;
Matches 335; Conservative 0; Mismatches 30; Indels 3; Gaps 1;

QY  2  AAGTATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTCT 61
Db   111 AGGTGCAACTTGTGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTCT 170
QY  62  CATGTGCAGCCTCTGATTACCTCAATAACTACGCCATGAATTCGGTCCGCGAGGCTC 121
Db   171 CATGTGCAGCCTCTGATTCAACTTCAATACCTACGCCATGAATTCGGTCCGCGAGGCTC 230
QY  122 CAGGAAGGGTTTGGAAATGGTTCGCGATAGAAGTAAAAGTAATAATTATGCAACAT 181
Db   231 CAGGAAGGGTTTGGAAATGGTTCGCGATAGAAGTAAAAGTAATAATTATGCAACAT 290
QY  182 TTTATCCGATTCAGTGAAGACAGGTTTACCATCTCCAGAGATGATTCACAAAGCATGC 241
Db   291 ATTATGCCGATTCATGAAGACAGATTCACCATCTCCAGAGATGATTCAGAAACATGC 350
QY  242 TCTATCTGCAATGAACAACCTTGAAACCTGAGGACACACCCATGTATTACTGTGTGAGAC 301
Db   351 TCTATCTGCAATGATCAACTTGAAAGCTGAGGACACACCCATGTATTACTGTGTGAGAC 410
QY  302 GGGGGGCTTCAGGATGTGCTATGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCT 361
Db   411 ---AGCATATAGTAACACTACGGTGCCTAGTACTACTGCTACTGCTACTGCTACTGCTACTGCT 467
QY  362 TCTCCTCA 369

```

```

|||||||
Db   468 TCTCCTCA 475

RESULT 4
MUSIGHAAAB
LOCUS      Mouse Ig heavy chain mRNA V-D-J region, 5' end.
DEFINITION M60021.1 GI:194548
ACCESSION  M60021
VERSION    M60021.1
KEYWORDS   C-region; D-region; J-region; V-region; immunoglobulin heavy chain.
SOURCE     Mouse (strain C57BL/6) spleen cell hybridoma 244, cDNA to mRNA.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
            1 (bases 1 to 499)
            Brigido.M.M. and Stollar,B.D.
            Two induced anti-z-DNA monoclonal antibodies use VH gene segments
            related to those of anti-DNA autoantibodies
            J. Immunol. 146 (6), 2005-2009 (1991)
            91170743
            PUBMED
            1900879
FEATURES
    source
        Location/Qualifiers
            1..499
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /map="chromosome 12"
                /cell_line="hybridoma 244"
                /tissue_type="liver"
                /dev_stage="fetus"
            1..499
                /gene="IgH"
            77..499
                /partial
            /gene="IgH"
            /codon_start=1
            /product="immunoglobulin heavy chain"
            /protein_id="AAAS1055.1"
            /db_xref="GI:194549"
            /translation="MLGLKVVFFVYQGVLCFVQLVESGGGLVQPKGSLKLSCAAS
            GTFNTYAHNVROAPGKLEWVARIRKSNVATYADSKDRFTISRDDSQSMLYL
            QMNLRTEDTAMYICVRQRLRYYYVNDYWGQGTSTVTVSS"
            77..133
                /gene="IgH"
            134..499
                /partial
            /gene="IgH"
            /product="immunoglobulin heavy chain"
            134..433
                /gene="IgH"
            434..448
                /gene="IgH"
            449..499
                /gene="IgH"
            /gene="IgH"
BASE COUNT    132 a    100 c    128 g    139 t
ORIGIN          Chromosome 12.
Query Match    82.2%; Score 303.4; DB 10; Length 499;
Best Local Similarity 90.8%; Pred. No. 8.6e-84;
Matches 335; Conservative 0; Mismatches 31; Indels 3; Gaps 1;

QY  1  GAAGTATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTC 60
Db   134 GAGTGCACGCTGTTGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGATCATTGAAACTC 193
QY  61  TCATGTGCAGCCTCTGGATTACCTTCAATACTACGCCATGAATTTGGGTCCGCCAGGCT 120
Db   194 TCATGTCCGCTCTGTTTACCTTCAATACCTATATGCCATGCATCGGGTCCGCCAGGCT 253
QY  121 CCAGGAAGGGTTTGAAGTGGGTTCGTCGCATAGAAGTAAAAGTAATAATTATGCAACA 180
Db   254 CCAGGAAGGGTTTGAAGTGGGTTCGTCGCATAGAAGTAAAAGTAATAATTATGCAACA 313

```



QY 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCCATGAATTTGGTCCGCCAGGCT 120  
Db 136 TCATGTGCAGCCTCTGGTTTCACTTCAATACCTATGCCATGAATGGTCCGCCAGGCT 195  
QY 121 CCAGGAAGGGTTTGAATGGGTTCCTCATAAAGTAAGTAAATATATGCAACA 180  
Db 196 CCAGGAAGGGTTTGAATGGGTTCCTCATAAAGTAAGTAAATATATGCAACA 255  
QY 181 TTTATGCGGATTCAGTGAAGACAGTTCCACATCTCCAGAGATGATTCACAAAGCATG 240  
Db 256 TATATGCGGATTCAGTGAAGACAGTTCCACATCTCCAGAGATGATTCACAAAGCATG 315  
QY 241 CTCATCTGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGATTACTGTGTGAGA 300  
Db 316 CTCATCTGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGATTACTGTGTGATC 375  
QY 301 CGGGGGGCTTCAGGATTCAGTATGCTATGGACTACTGGGTCAAGGAACCTCACTACCC 360  
Db 376 CTCATGCGTAACTACGCCCTCTATGCTATGGACTACTGGGTCAAGGAACCTCACTACCC 435  
QY 361 GTCTCCTCA 369  
Db 436 GTCTCCTCA 444

RESULT 7  
LOCUS I73737 357 bp DNA linear PAT 03-APR-1998  
DEFINITION Sequence 1 from patent US 5686600.  
ACCESSION I73737  
VERSION I73737.1 GI:3009878  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 357)  
AUTHORS Carozzi,N.B. and Koziel,M.G.  
TITLE Antibodies which bind to insect gut proteins and their use  
JOURNAL Patent: US 5686600-A 11-NOV-1997;  
FEATURES Location/Qualifiers  
source 1..357  
BASE COUNT 102 a 76 c 91 g 88 t  
ORIGIN

Query Match 81.4%; Score 300.4; DB 6; Length 357;  
Best Local Similarity 91.0%; Pred. No. 7.3e-83;  
Matches 335; Conservative 0; Mismatches 21; Indels 12; Gaps 1;  
QY 2 AAGTGTGCTGGTGGAGTCTGGTGGAGGATTTGGTGCAGCCTAAAGGTCATTGAAACTCT 61  
Db 2 AGTCAAACTGCAGGAGTCTGGTGGAGGATTTGGTGCAGCCTAAAGGTCATTGAAACTCT 61  
QY 62 CATGTGCAGCCTCTGGATTACCTTCAATAAAGTAAGTAAATATATGCAACAT 181  
Db 62 CATGTGCAGCCTCTGGATTACCTTCAATAAAGTAAGTAAATATATGCAACAT 181  
QY 122 CAGGAAGGGTTTGAATGGGTTCCTCATAAAGTAAGTAAATATATGCAACAT 241  
Db 122 CAGGAAGGGTTTGAATGGGTTCCTCATAAAGTAAGTAAATATATGCAACAT 241  
QY 182 TTTATGCGGATTCAGTGAAGACAGTTCCACATCTCCAGAGATGATTCACAAAGCATG 301  
Db 182 TTTATGCGGATTCAGTGAAGACAGTTCCACATCTCCAGAGATGATTCACAAAGCATG 301  
QY 242 TCATCTGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGATTACTGTGTGAGAC 361  
Db 242 TCATCTGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGATTACTGTGTGAGAC 361  
QY 302 GGGGGGCTTCAGGATTCAGTATGCTATGGACTACTGGGTCAAGGAACCTCACTACCCG 435  
Db 302 GGGGGGCTTCAGGATTCAGTATGCTATGGACTACTGGGTCAAGGAACCTCACTACCCG 435  
QY 361 GTCTCCTCA 369  
Db 436 GTCTCCTCA 444

QY 362 TCTCCTCA 369  
Db 350 TCTCCTCA 357

RESULT 8  
LOCUS I73745 1797 bp DNA linear PAT 03-APR-1998  
DEFINITION Sequence 17 from patent US 5686600.  
ACCESSION I73745  
VERSION I73745.1 GI:3009886  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1797)  
AUTHORS Carozzi,N.B. and Koziel,M.G.  
TITLE Antibodies which bind to insect gut proteins and their use  
JOURNAL Patent: US 5686600-A 17-11-NOV-1997;  
FEATURES Location/Qualifiers  
source 1..1797  
BASE COUNT 455 a 488 c 452 g 402 t  
ORIGIN

Query Match 81.4%; Score 300.4; DB 6; Length 1797;  
Best Local Similarity 91.0%; Pred. No. 8e-83;  
Matches 335; Conservative 0; Mismatches 21; Indels 12; Gaps 1;  
QY 2 AAGTGTGCTGGTGGAGTCTGGTGGAGGATTTGGTGCAGCCTAAAGGTCATTGAAACTCT 61  
Db 464 AGGTCAAACTGCAGGAGTCTGGTGGAGGATTTGGTGCAGCCTAAAGGTCATTGAAACTCT 523  
QY 62 CATGTGCAGCCTCTGGATTACCTTCAATAAAGTAAGTAAATATATGCAACAT 181  
Db 524 CATGTGCAGCCTCTGGATTACCTTCAATAAAGTAAGTAAATATATGCAACAT 181  
QY 122 CAGGAAGGGTTTGAATGGGTTCCTCATAAAGTAAGTAAATATATGCAACAT 241  
Db 584 CAGGAAGGGTTTGAATGGGTTCCTCATAAAGTAAGTAAATATATGCAACAT 241  
QY 182 TTTATGCGGATTCAGTGAAGACAGTTCCACATCTCCAGAGATGATTCACAAAGCATG 301  
Db 644 CTTATGGCGATTCAGTGAAGACAGTTCCACCTCTCCAGAGATGATTCACAAAGCATG 703  
QY 242 TCATCTGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGATTACTGTGTGAGAC 361  
Db 704 TCATCTGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGATTACTGTGTGAGAC 361  
QY 302 GGGGGGCTTCAGGATTCAGTATGCTATGGACTACTGGGTCAAGGAACCTCACTACCCG 811  
Db 762 -----GGTAGTATACGGTGTCTATGGACTACTGGGTCAAGGAACCTCACTACCCG 811  
QY 362 TCTCCTCA 369  
Db 812 TCTCCTCA 819

RESULT 9  
LOCUS AF178591 366 bp mRNA linear ROD 22-MAY-2000  
DEFINITION Mus musculus 23-7 immunoglobulin heavy chain variable region mRNA, partial cds.  
ACCESSION AF178591  
VERSION AF178591.1 GI:5853171  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 366)  
AUTHORS Puterman,C., Deocharan,B. and Diamond,B.  
TITLE Molecular analysis of the autoantibody response in peptide-induced



```

KEYWORDS
SOURCE      Mus musculus
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 366)
AUTHORS    Putterman,C., Deocharan,B. and Diamond,B.
TITLE      Molecular analysis of the autoantibody response in peptide-induced
           autoimmunity
JOURNAL    J. Immunol. 164 (5), 2542-2549 (2000)
MEDLINE    20143847
PUBMED     10679092
REFERENCE   2 (bases 1 to 366)
AUTHORS    Putterman,C., Deocharan,B. and Diamond,B.
TITLE      Direct Submission
JOURNAL    Submitted (17-AUG-1999) Rheumatology/Medicine, Albert Einstein
           College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA
FEATURES   source
           1..366
           /organism="Mus musculus"
           /strain="BALB/c"
           /db_xref="taxon:10090"
           /cell_line="39-9"
           /note="from peptide-immunized mouse"
           <1..>366
           /note="anti-peptide antibody"
           /codon_start=1
           /product="immunoglobulin heavy chain variable region"
           /protein_id="AAD54352.1"
           /db_xref="GI:5853182"
           /translation="DKVSVESGGGLVOPKSLKLSCAASGFTNTYAMNVROAPGKG
           LEWARIKRSNYATYYADSVKDRFTISRDSQSMYLVQNNLKTEDTAMYCYVRIG
           LGRIYEDVWGAGTFTVTSS"
BASE COUNT 100 a 82 c 98 g 86 t
ORIGIN
Query Match      80.5%; Score 297; DB 10; Length 366;
Best Local Similarity 89.7%; Pred. No. 8.4e-82;
Matches 331; Conservative 0; Mismatches 35; Indels 3; Gaps 1;

Qy 1 GAAGTGATGCTGGTGAGCTCTGGTGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GAGGTCAAGGTGGTGAGCTCTGGAGGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 60

Qy 61 TCATGTGAGCCTCTGGATTACCTTCAATCACTACGCGCATGAATGGTCCGCCAGGCT 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TCATGTGAGCCTCTGGATTACCTTCAATCACTACGCGCATGAATGGTCCGCCAGGCT 120

Qy 121 CCAGAAAGGTTTGGCAATGGTCTCGCATAGCAAGTAAAGTAATAATTATGCAACA 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CCAGAAAGGTTTGGCAATGGTCTCGCATAGCAAGTAAAGTAATAATTATGCAACA 180

Qy 181 TTTTATGCCGATTCAGTGAAGACAGGTTACCACTCTCCAGAGATGATTCACAAAGCATG 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TATTATGCCGATTCAGTGAAGACAGGTTACCACTCTCCAGAGATGATTCACAAAGCATG 240

Qy 241 CTCATCTGCAATGAACAACTTGAAGTGAAGGACAGGACGATGATGATGATGATGATG 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CTCATCTGCAATGAACAACTTGAAGTGAAGGACAGGACGATGATGATGATGATGATG 300

Qy 301 CGGGGGGCTTACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 CATGGG---CGGGGGCGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCTACT 357

Qy 361 GTCTCCTCA 369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 GTCTCCTCA 366

RESULT 12
AF045483
LOCUS      AF045483          408 bp    mRNA    linear    ROD 24-SEP-2001

DEFINITION Mus musculus dCl anti-poly(dC) monoclonal antibody heavy chain
           variable region, (Igh) mRNA, partial cds.
ACCESSION AF045483
VERSION   AF045483.1 GI:2906049
KEYWORDS Mus musculus.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 408)
AUTHORS Li,J., Fernandez,L., O'Connor,K.C., Imanishi-Kari,T. and
           Stollar,B.D.
TITLE      The rearranged V(H) domain of a physiologically selected
           anti-single-stranded DNA antibody as a precursor for formation of
           IgM and IgG antibodies to diverse antigens
JOURNAL    J. Immunol. 167 (7), 3746-3755 (2001)
MEDLINE    21448704
PUBMED     11564791
REFERENCE 2 (bases 1 to 408)
AUTHORS O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
TITLE      Anti-DNA antibodies of normal mice immunized with poly(dC) are
           structurally similar to natural autoantibodies
JOURNAL    Unpublished
REFERENCE 3 (bases 1 to 408)
AUTHORS O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
TITLE      Direct Submission
JOURNAL    Submitted (02-FEB-1998) Biochemistry, Tufts University School of
           Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
FEATURES   source
           1..408
           /organism="Mus musculus"
           /strain="C57BL/6"
           /db_xref="taxon:10090"
           /chromosome="12"
           /clone="dCl"
           /cell_line="spleen cell hybridoma"
           <1..>408
           /gene="Igh"
           1..>408
           /gene="Igh"
           /codon_start=1
           /product="anti-poly(dC) monoclonal antibody heavy chain"
           /protein_id="AAC04511.1"
           /db_xref="GI:2906050"
           /translation="MVLGLKWFVFFVQVHCFLVQVSGGLVOPKGSLSKLSCAAS
           GFPTNTYAMHVRQAPKGLKLEWARIKRSNYATYYADSVKDRFTISRDSQSMYLV
           QNNLKTEDTAMYCYVRSDAMDYWGQTSVTSS"
           1..57
           /gene="Igh"
           58..357
           /gene="Igh"
           358..375
           /gene="Igh"
           376..408
           /gene="Igh"
BASE COUNT 105 a 83 c 104 g 116 t
ORIGIN
Query Match      78.9%; Score 291; DB 10; Length 408;
Best Local Similarity 89.7%; Pred. No. 6.4e-80;
Matches 331; Conservative 0; Mismatches 20; Indels 18; Gaps 1;

Qy 1 GAAGTGATGCTGGTGAGCTCTGGTGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 GAGGTGCAGCTCTGTGAGTCTGGTGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 117

Qy 61 TCATGTGAGCCTCTGGATTACCTTCAATCACTACGCGCATGAATGGTCCGCCAGGCT 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 TCATGTGAGCCTCTGGTTTACCTTCAATCACTATGCATGCACCTGGTCCGCCAGGCT 177

Qy 121 CCAGAAAGGTTTGGCAATGGTCTCGCATAGCAAGTAAAGTAATAATTATGCAACA 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 CCAGAAAGGTTTGGCAATGGTCTCGCATAGCAAGTAAAGTAATAATTATGCAACA 237

```

Qy	181	TTTTATGCCGANTCAGTGAAGACAGAGTTCACCATCTCCAGAGATGATTCACAAGCATG	240
I			
Db	238	TATTATGCCGATTTCAGTGAAGACAGATTCACCATCTCCAGAGATGATTCACAAGCATG	297
I			
Qy	241	CTCTATCTGCAAAATGACAACTTCAAAACTCAGACACACGCCATGTATTACTGTGTGAGA	300
I			
Db	298	CTCTATCTGCAAAATGACAACTTCAAAACTCAGACACACGCCATGTATTACTGTGTGAGA	357
I			
Qy	301	CGGGGGGCTTCAGGAGTTGACTGCTATGGACTACTTGGGTCGAAGAACTCACTCACTCAC	360
I			
Db	358	-----TCCGATGCTATGACTACTTGGGTCGAAGAACTCACTCACTCAC	399
Qy	361	GTCCTCTCA	369
I			
Db	400	GTCCTCTCA	408
RESULT	13		
MUSIGH125A			
LOCUS	MUSIGH125A	366 bp	mrna
DEFINITION	Mus musculus anticytotoxic Ig 1255 H-chain mRNA V region,		linear ROD 05-APR-1994
	partial cds.		
ACCESSION	L31512		
VERSION	L31512.1	GI:468344	
KEYWORDS	V region; immunoglobulin heavy chain.		
SOURCE	Mus musculus (strain BALB/c, sub_species domesticus) hybridoma		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 366)		
	Casadevall, A., Deshaw, M., Fan, M., Dromer, F., Kozel, T.R. and		
	Pirofski, L.-A.		
TITLE	Molecular and idiotypic analysis of antibodies to Cryptococcus		
	neoformans		
JOURNAL	Unpublished (1994)		
FEATURES	Location/Qualifiers		
source	1..366		
	/organism="Mus musculus"		
	/strain="BALB/c"		
	/sub_species="domesticus"		
	/db_xref="taxon:10090"		
	/cell_line="1255"		
	/cell_type="B-cell"		
	/tissue_type="hybridoma"		
	<1..>366		
CDS	/codon_start=1		
	/product="immunoglobulin heavy chain V region"		
	/protein_id="AA37928.1"		
	/db_xref="GI:468345"		
	/translation="EVOLVETGGVLQPKSLKVCATSGTFTNTVMVSWROAPKGL		
	LEWRAIKNKINNYGYADSVKDRFTISRDDSHNLIYLOMSNLKTEDTAMYYCYCRDRL		
	RGWDYAMAYWGQTSVTSS"		
BASE COUNT	104 a	75 c	93 t
ORIGIN			1 others
Query Match	78.5%	Score 289.6;	DB 10;
Best Local Similarity	88.3%;	Pred. No. 1.7e-79;	Length 366;
Matches	326;	Conservative 0;	Mismatches 40;
			Indels 3; Gaps 1;
Qy	1	GAAGTGATGCTGGTGGAGTCTGTTGGAGGATTTGGTGAGCCCTAAAGGTCATTGAAACTC	60
I			
Db	1	GAGGTGCACCTTGTGAGACTGGTGGAGGATTTGGTGAGCCCTAAAGGTCATTGAAACTC	60
I			
Qy	61	TCATGTGCAGCCTCTGGATTACCTTCAATAACCTACGCCATGAATGGGTCGCCGACGCT	120
I			
Db	61	TCATGTGCAACTTCTGGATTACCTTCAATAACCAATGTCATGAGCTGGTCCGCCAGGCT	120
I			
Qy	121	CCAGAAAGGGTTGGAAATGGGTTGCTCGCATAGAGATAAGAACTAATATTGCAACA	180
I			
Db	121	CCAGAAAGGGTTGGAAATGGGTTGCTCGCATAAATAAATAATTAATATTGGAACA	180
I			

Qy	181	TTTTATCCGGGATTCAGTGAAGACAGAGGTTCCACCATCTCCAGAGATGATTTCACAAAGCATG	240
Db	181	TATTTATCCGGATTCAGTGAAGACAGAGGTTCCACCATCTCCAGAGATGATTTCACAAACATT	240
Qy	241	CTCTATCTGCAATGAACAACCTTCAAAACTGAGGACACAGGCATGTATTACTGTGTGAGA	300
Db	241	CTCTATCTGCAATGAGCAATTTGAAACTGAGGACACAGGCATGTATTACTGTGTGAGA	300
Qy	301	CGGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGTCAANGAACCTCACTCACC	360
Db	301	--CATCTGAGGGCTGGGACTATGCTATGGCTTGGGCTCAGGGAACCTCAGTCACC	357
Qy	361	GTCTCCTCA	369
Db	358	GTCTCCTCA	366
RESULT 14			
MUSHVCRA			
LOCUS			
DEFINITION			
MUSHVCRA			
Mouse mRNA for immunoglobulin heavy chain variable region, partial			
sequence.			
ACCESSION			
D50376			
VERSION			
D50376.1 GI:804892			
KEYWORDS			
Immunoglobulin heavy chain variable region.			
SOURCE			
Mus musculus (strain BALB/c) hybridoma producing anti-human			
procollagenase antibody, cell-line K1E6 cDNA to mRNA.			
ORGANISM			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
1 (bases 1 to 360)			
Yamanaka,H.I., Kirii,Y. and Ohmoto,H.			
An improved phage display antibody cloning system using newly			
designed PCR primers optimized for pfu DNA polymerase			
J. Biochem. 117 (6), 1218-1227 (1995)			
JOURNAL			
MEDLINE			
96104992			
REFERENCE			
2 (bases 1 to 360)			
Yamanaka,H.			
Direct Submission			
AUTHORS			
TITLE			
Submitted (21-APR-1995) Hachiro Yamanaka, New Drug Research			
Laboratories, Kanebo Ltd.; 5-90 Tomobuchi-cho 1-chome,			
Miyakojima-ku, Osaka, Osaka 534, Japan (Tel:06-921-1281(ex.3609),			
Fax:06-923-3381)			
COMMENT			
Submitted (21-Apr-1995) to DBJ by:			
Hachiro Yamanaka			
New Drug Research Laboratories.			
Kanebo Ltd.			
1-5-90 Tomobuchi-cho Miyakojima-ku			
Osaka, Osaka 534			
Japan			
Phone: 06-921-1281 x3609			
Fax: 06-923-3381.			
FEATURES			
source			
Location/Qualifiers			
1. .360			
/organism="Mus musculus"			
/strain="BALB/c"			
/sub_species="domesticus"			
/db_xref="taxon:10090"			
/cell_line="K1E6"			
/cell_type="hybridoma producing anti-human procollagenase			
antibody"			
<1. .>360			
V_region			
/product="immunoglobulin heavy chain variable region"			
/db_xref="IMGT:D50376"			
primer_bind			
1. .26			
/note="PCR primer"			
primer_bind			
complement(339. .360)			
/note="PCR primer"			
BASE COUNT			
10			





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:34:25 ; Search time 198.734 Seconds  
(without alignments)  
4181.413 Million cell updates/sec

Title: US-09-893-615-86

Perfect score: 369

Sequence: 1 GAAGTGATGCTGGTGAGTC.....CCTCACTACCGTGCTCTCA 369

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	300.4	81.4	357	17 AAT15725	3B1 heavy chain va
2	300.4	81.4	1797	17 AAT15733	3B1 single chain a
3	293.8	79.6	1079	21 AAT51115	Anti-Her2neu singl
4	288.6	78.2	345	18 AAT58267	Lead binding MAb 2
5	284.4	77.1	330	18 AAT58261	Lead binding MAB 6
6	283.8	76.9	375	21 AAA40202	H. pylori 26 kda p
7	269.2	73.0	443	22 AAD13178	Murine antibody 1D
8	249.2	67.5	357	22 AAD13179	Humanised murine a
9	227.2	61.6	366	21 AAA38896	2G3 hybridoma VH d

10	225.8	61.2	359	21	AAA44338	Human secreted exp
11	215.4	58.4	5227	18	AAT79537	Plasmid pTERMSC2H1
12	214.6	58.2	369	22	AAS03515	DNA encoding anti-
13	214.6	58.2	761	24	ABK47271	Insulin/insulin-11
14	213.6	57.9	372	18	AAT58268	Lead binding MAB 2
15	212.8	57.7	768	24	AAS97139	P5-2 single chain
16	212.8	57.7	1509	24	AAS97147	3B10xP5-2 bispecif
17	212.2	57.5	375	24	ABK88128	DNA encoding mouse
18	210.6	57.1	761	24	ABK47275	Insulin/insulin-11
19	210.2	57.0	357	19	AAF22331	Nucleic acid encod
20	209	56.6	378	22	AAF29062	Human HIV-1 monoc
21	205	55.6	361	13	AAQ29148	WN1 222-5 antibody
22	204.4	55.4	379	22	AAI68772	Human autoantibody
23	204.2	55.3	738	14	AAQ43293	Sequence encoding
24	204.2	55.3	738	15	AAQ66841	CC49 VL / 217 / 4-
25	204.2	55.3	738	20	AAQ4747	DNA encoding a pro
26	204.2	55.3	738	20	AAV99764	Fusion polypeptide
27	204.2	55.3	738	21	AAV95090	DNA encoding bival
28	204.2	55.3	738	21	AAA59622	DNA encoding a sin
29	204.2	55.3	738	21	AA257095	4-4-20 Vh region a
30	204.2	55.3	738	21	AA290344	DNA (SEQ ID NO:22)
31	204.2	55.3	738	21	AA237393	Antibody CC49/anti
32	204.2	55.3	744	14	AAQ43288	Sequence encoding
33	204.2	55.3	744	20	AAQ4742	DNA encoding a pro
34	204.2	55.3	744	21	AAA95085	Bivalent antigen b
35	204.2	55.3	744	21	AAA59617	DNA encoding a het
36	204.2	55.3	744	21	AA257090	4-4-20 Vh region a
37	204.2	55.3	744	21	AA290339	DNA (SEQ ID NO:12)
38	204.2	55.3	752	24	ABK47272	Insulin/insulin-11
39	203	55.0	345	14	AAQ40046	BW 835 Vh. Synthe
40	202.6	54.9	1172	20	AAH87209	AGA2-HA-4-4-20 gen
41	202.6	54.9	1172	22	AAH87209	AGA2-HA-4-4-20 gen
42	202.6	54.9	1172	24	AAH27164	Yeast AGA2 aggluti
43	202.2	54.8	357	22	AAS03439	Yeast cell wall pr
44	202.2	54.8	360	22	AAH47724	DNA encoding anti-
45	201.8	54.7	1774	20	AAZ20419	Nucleotide sequenc
						Antibody ABX-CBL h

#### ALIGNMENTS

#### RESULT 1

AAT15725  
ID AAT15725 standard; cDNA; 357 BP.

XX AAT15725;

XX AC AAT15725;

XX DT 25-JUN-1996 (first entry)

XX DE 3B1 heavy chain variable region DNA from pcIB4613.

XX KW delta endotoxin; Bacillus thuringiensis; western corn rootworm;

XX KW WCRW; maize; pesticide; brush border membrane vesicle; monoclonal;

XX KW antibody; ss.

XX OS Insecta sp.

XX PN WO9600783-A1.

XX XX

XX PD 11-JAN-1996.

XX XX

XX PF 20-JUN-1995; 95WO-IB00497.

XX PR 28-JUN-1994; 94US-0267641.

XX XX

XX PA (CIBA ) CIBA GEIGY AG.

XX XX

XX PI Carozzi NB, Koziel MG;

XX XX

XX PI WPI; 1996-077494/08.

XX DR P-PSDB; AAR90829.

XX XX

XX PT New monoclonal antibodies which bind insect gut proteins - used

PT partic. with toxin moieties for the control of insect pests, partic.  
 XX in plants  
 PS Claim 8; Page 50-51; 106pp; English.  
 XX  
 CC AAT15725-35 are DNA mols. encoding a monoclonal antibody or a binding  
 CC fragment. Mabs were produced by using insect guts, partic. insect  
 CC brush border membranes (BBMs), esp. corn rootworm, as antigen;  
 CC immunising a donor animal with the antigen; isolating immunocompetent B  
 CC cells from the immunised animal; fusing B cells with a tumour cell line;  
 CC isolating the fused cells, culturing them and cloning positive hybrid  
 CC cells; and screening the hybrid cells for prodn. of the required MABs.  
 CC The Mabs bind to the gut of a target insect but do not bind to mammalian  
 CC BBMs. The DNA sequence can be operably linked to a toxin moiety, esp.  
 CC etc... The Abs are useful for control of insect pests, e.g. Coleoptera,  
 CC Diptera, Hymenoptera and Lepidoptera. The pesticidal compsn. is pref.  
 CC applied to a plant, e.g. maize.  
 XX  
 SQ Sequence 357 BP; 102 A; 76 C; 91 G; 88 T; 0 other;  
 Query Match 81.4%; Score 300.4; DB 17; Length 357;  
 Best Local Similarity 91.0%; Pred. No. 3.9e-80;  
 Matches 335; Conservative 0; Mismatches 21; Indels 12; Gaps 1;  
 QY 2 AAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGAGGCTAAAGGCTCATTTGAAACTCT 61  
 DB 2 AGGTCAAACTGCAGGAGTCTGGTGGAGGATTGGTGAGGCTAAAGGCTCATTTGAAACTCT 61  
 QY 62 CATGTGAGCCTCTGGATTCACCTTCAATAAATTCGCCATGAATGGGTCGCGCAGGCTC 121  
 DB 62 CATGTGAGCCTCTGGATTCACCTTCAATAAATTCGCCATGAATGGGTCGCGCAGGCTC 121  
 QY 122 CAGGAAGGGTTTGGATGGTGGTCTCGCATAGAGTAAAGTAAATATATGCAACAT 181  
 DB 122 CAGGAAGGGTTTGGATGGTGGTCTCGCATAGAGTAAAGTAAATATATGCAACAT 181  
 QY 182 TTTATCCGATTCAGTGAAGACAGAGTTCACCATCTCCAGAGATGATTCACAAGCATGC 241  
 DB 182 CTATGCGGATTCAGTGAAGACAGAGTTCACCGTCTCCAGAGATGATTCACAAGCATGT 241  
 QY 242 TCTATCTGCAATGAACAACCTTGAACACTGAGGACACAGCCATGATTAATGTTGAGAC 301  
 DB 242 TCTATCTGCAATGAACAACCTTGAACACTGAGGACACAGCCATGATTAATGTTGAG-- 299  
 QY 302 GGGGGGCTTCAGGGATTGACTATGCTATGAGTACTGCGTCAAGGAACCTCACTCACCG 361  
 DB 300 -----GGTAGTATACGGTCTATGAGTACTGCGTCAAGGAACCTCACTCACCG 349  
 QY 362 TCTCCTCA 369  
 DB 350 TCTCCTCA 357  
 RESULT 2  
 AAT15733  
 ID AAT15733 standard; cdna; 1797 BP.  
 XX  
 AC AAT15733;  
 XX  
 XX  
 DT 25-JUN-1996 (first entry)  
 XX  
 DE 3B1 single chain antibody DNA from pCIB4631.  
 XX  
 KW delta endotoxin; Bacillus thuringiensis; western corn rootworm;  
 KW WCRW; maize; pesticide; brush border membrane vesicle; monoclonal;  
 KW antibody; ss.  
 XX  
 OS Insecta sp.  
 XX  
 PN WO9600783-A1.  
 XX  
 PD 11-JAN-1996.

XX  
 PF 20-JUN-1995; 95WO-IB00497.  
 XX  
 PR 28-JUN-1994; 94US-0267641.  
 XX  
 PA (CIBA ) CIBA GEIGY AG.  
 XX  
 PI Carozzi NB, Koziel MG;  
 XX  
 XX WPI: 1996-077494/08.  
 DR P-PSDB; AAR90837.  
 XX  
 XX New monoclonal antibodies which bind insect gut proteins - used  
 PT partic. with toxin moieties for the control of insect pests, partic.  
 PT in plants  
 XX  
 PS Claim 8; Page 68-72; 106pp; English.  
 XX  
 CC AAT15725-35 are DNA mols. encoding a monoclonal antibody or a binding  
 CC fragment. Mabs were produced by using insect guts, partic. insect  
 CC brush border membranes (BBMs), esp. corn rootworm, as antigen;  
 CC immunising a donor animal with the antigen; isolating immunocompetent B  
 CC cells from the immunised animal; fusing B cells with a tumour cell line;  
 CC isolating the fused cells, culturing them and cloning positive hybrid  
 CC cells; and screening the hybrid cells for prodn. of the required MABs.  
 CC The Mabs bind to the gut of a target insect but do not bind to mammalian  
 CC BBMs. The DNA sequence can be operably linked to a toxin moiety, esp.  
 CC selected from e.g. Bacillus toxins, Pseudomonas exotoxin and phytolaccin,  
 CC etc... The Abs are useful for control of insect pests, e.g. Coleoptera,  
 CC Diptera, Hymenoptera and Lepidoptera. The pesticidal compsn. is pref.  
 CC applied to a plant, e.g. maize.  
 XX  
 SQ Sequence 1797 BP; 455 A; 488 C; 452 G; 402 T; 0 other;  
 Query Match 81.4%; Score 300.4; DB 17; Length 1797;  
 Best Local Similarity 91.0%; Pred. No. 6.8e-80;  
 Matches 335; Conservative 0; Mismatches 21; Indels 12; Gaps 1;  
 QY 2 AAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGAGGCTAAAGGCTCATTTGAAACTCT 61  
 DB 464 AGGTCAAACTGCAGGAGTCTGGTGGAGGATTGGTGAGGCTAAAGGCTCATTTGAAACTCT 523  
 QY 62 CATGTGAGCCTCTGGATTCACCTTCAATAAATTCGCCATGAATGGGTCGCGCAGGCTC 121  
 DB 524 CATGTGAGCCTCTGGATTCACCTTCAATAAATTCGCCATGAATGGGTCGCGCAGGCTC 583  
 QY 122 CAGGAAGGGTTTGGATGGTGGTCTCGCATAGAGTAAAGTAAATATATGCAACAT 181  
 DB 584 CAGGAAGGGTTTGGATGGTGGTCTCGCATAGAGTAAAGTAAATATATGCAACAT 643  
 QY 182 TTTATCCGATTCAGTGAAGACAGAGTTCACCATCTCCAGAGATGATTCACAAGCATGC 241  
 DB 644 CTATGCGGATTCAGTGAAGACAGAGTTCACCGTCTCCAGAGATGATTCACAAGCATGT 703  
 QY 242 TCTATCTGCAATGAACAACCTTGAACACTGAGGACACAGCCATGATTAATGTTGAGAC 301  
 DB 704 TCTATCTGCAATGAACAACCTTGAACACTGAGGACACAGCCATGATTAATGTTGAG-- 761  
 QY 302 GGGGGGCTTCAGGGATTGACTATGCTATGAGTACTGCGTCAAGGAACCTCACTCACCG 361  
 DB 762 -----GGTAGTATACGGTCTATGAGTACTGCGTCAAGGAACCTCACTCACCG 811  
 QY 362 TCTCCTCA 369  
 DB 812 TCTCCTCA 819  
 RESULT 3  
 AAZ51115  
 ID AAZ51115 standard; DNA; 1079 BP.  
 XX  
 AC AAZ51115;  
 XX







CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is a DNA encoding murine antibody 1D9 heavy chain  
 CC variable region.

XX SQ Sequence 443 BP; 117 A; 97 C; 108 G; 121 T; 0 other;

Query Match 73.0%; Score 269.2; DB 22; Length 443;  
 Best Local Similarity 94.0%; Pred. No. 9.4e-71;  
 Matches 280; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 60  
 DB 58 GAGGTGCAGCTTGTAGCTGGAGGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 117  
 QY 61 TCATGTGCAGCCTCTGGATTACCTTCAATAACTACGCCATGAATTTGGTCCGCGAGGCT 120  
 DB 118 TCATGTGCAGCCTCTGGATTACCTTCAATCCCTACGCCATGAATTTGGTCCGCGAGGCT 177  
 QY 121 CCAGGAAGGGTTTGGATGGTCTGCGCATAGAGTAAAGTAAATATTTATGCAACA 180  
 DB 178 CCAGGAAGGGTTTGGATGGTCTGCGCATAGAGTAAATATTTATGCAACA 237  
 QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTTCCATCTCCAGAGATGATTCACAAAGCATG 240  
 DB 238 TATTATGCCGATTCAGTGAAGACAGATACACCATCTCCAGAGATGATTCAGAAAGTATG 297  
 QY 241 CTCATCTGCAATGACAACTTGAACACTGAGGACAGCCATCTATTCTGTGTGA 298  
 DB 298 CTCATCTGCAATGACAACTTGAACACTGAGGACAGCCATCTATTCTGTGTGA 355

# RESULT 8

AAD13179

ID AAD13179 standard; DNA; 357 BP.

XX AC AAD13179;

XX DT 16-OCT-2001 (first entry)

XX DE Humanised murine antibody heavy chain 1D9RHA DNA.

XX KW Murine; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KW inflammatory disorder; autoimmune disease; rheumatoid arthritis; shock;  
 KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention;  
 KW neointimal hyperplasia; antibody 1D9 heavy chain; 1D9RHA; ds.

XX OS Chimeric - Mus sp.

OS OS Chimeric - Homo sapiens.

XX PH Key Location/Qualifiers

FT CDS 1..357

FT /\*tag= a

FT /product= "Humanised murine antibody heavy chain

FT region, 1D9RHA"

FT /note= "CDS does not include start and stop codon"

FT /partial

XX PN W0200157226-AL.

XX PN

PD 09-AUG-2001.

XX 02-FEB-2001; 2001WO-US03537.

XX 03-FEB-2000; 2000US-0497625.

XX (MILL-) MILLENNIUM PHARM INC.

PI Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;

XX WPI; 2001-488888/53.

DR P-PSDB; AAE07034.

XX Humanized immunoglobulin for treating a CC-chemokine receptor

PT 2-mediated disorder in a patient, comprises a binding specificity for

PT CCR2, and a non-human antigen binding region and human immunoglobulin

PT -

XX Claim 64; Fig 23; 183pp; English.

XX The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is a DNA encoding humanised murine antibody heavy  
 CC chain region, 1D9RHA. 1D9RHA sequence consist of the complementarity  
 CC determining regions (CDRs) of the murine 1D9 antibody heavy chain  
 CC variable (VH) region genetically inserted into the framework regions  
 CC (FRs) of the human 4B4'CL antibody VH region.

XX Sequence 357 BP; 99 A; 79 C; 95 G; 84 T; 0 other;

QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 60  
 DB 1 GAGGTGCAGCTTGTAGCTGGAGGAGGATTGGTGCAGCCTAAAGGTCATTGAAACTC 60  
 QY 61 TCATGTGCAGCCTCTGGATTACCTTCAATAACTACGCCATGAATTTGGTCCGCGAGGCT 120  
 DB 61 TCATGTGCAGCCTCTGGATTACCTTCAATAACTACGCCATGAATTTGGTCCGCGAGGCT 120  
 QY 121 CCAGGAAGGGTTTGGATGGTCTGCGCATAGAGTAAAGTAAATATTTATGCAACA 180  
 DB 121 CCAGGAAGGGTTTGGATGGTCTGCGCATAGAGTAAATATTTATGCAACA 180  
 QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTTCCATCTCCAGAGATGATTCACAAAGCATG 240  
 DB 181 TATTATGCCGATTCAGTGAAGACAGATACACCATCTCCAGAGATGATTCACAAAGCATG 240  
 QY 241 CTCATCTGCAATGACAACTTGAACACTGAGGACAGCCATCTATTCTGTGT 294  
 DB 241 CTCATCTGCAATGACAACTTGAACACTGAGGACAGCCATCTATTCTGTGT 294

Query Match 67.5%; Score 249.2; DB 22; Length 357;

Best Local Similarity 90.5%; Pred. No. 8.6e-65;

Matches 266; Conservative 0; Mismatches 28; Indels 0; Gaps 0;



```
RESULT 9
AAA38896
ID AAA38896 standard; DNA; 366 BP.
XX
AC AAA38896;
XX
DT 29-AUG-2000 (first entry)
XX
DE 2G3 hybridoma VH domain encoding DNA SEQ ID NO:1.
XX
KW Antigen binding site; immunoglobulin; cancer antigen; immunological;
KW antibody; tumour; human; mucin; cancer; cytostatic; hybridoma;
KW specific binding assay; affinity purification; drug targeting;
KW toxin targeting; imaging; genetic; therapeutic; ss.
XX
OS Homo sapiens.
XX
PN US6054561-A.
XX
PD 25-APR-2000.
XX
PF 07-JUN-1995; 95US-0483749.
XX
PR 21-MAR-1986; 86US-0842476.
PR 08-MAY-1988; 88US-0190778.
PR 08-FEB-1984; 84US-0577976.
PR 11-JAN-1985; 85US-0690750.
PR 11-AUG-1994; 94US-0288981.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Ring DB;
XX
DR WPI: 2000-338508/29.
DR P-PSDB; AAY90812.
XX
PT Monoclonal antibody capable of binding to human breast cancer antigen
PT useful for affinity purification, drug or toxin targeting, imaging, and
PT treating cancer
XX
PS Disclosure; Fig 1; 57pp; English.
XX
CC The present invention describes a monoclonal antibody (MAB) (I) that
CC binds to a human breast cancer antigen that is also bound by MAB 454C11
CC and 520C9 (produced hybridoma ATCC HB8484 and HB8696, respectively). Also
CC described is a hybridoma that produces (I). (I) is useful in specific
CC binding assays, affinity purification, drug or toxin targeting, imaging,
CC and genetic or immunological therapeutics for various cancers. The
CC present sequence encodes a VH domain derived from a 2G3 hybridoma, which
CC is used in the exemplification of the present invention.
XX
SQ Sequence 366 BP; 103 A; 73 C; 95 G; 95 T; 0 other;

Query Match 61.6%; Score 227.2; DB 21; Length 366;
Best Local Similarity 78.0%; Pred. No. 3.4e-58;
Matches 287; Conservative 0; Mismatches 78; Indels 3; Gaps 1;

QY 1 GAAGTCATGCTGGTGAGTCTGGTGAGGATGGTGACCCCTAAAGGGTCATTGAACATC 60
DB 1 GAAGTGAAGCTTGAGAGGCTGGAGAGGCTGGTGCAACCTGGAAGATCCATGAACATC 60
QY 61 TCATGTGAGCCCTCTGGATTACCTTCAATACTAGCCATGAATGGGTCCGCCAGGCT 120
DB 61 TCTGTGTGCTTCCCTCGATCTGATTTCACTAGTAAGTATGATTTGAATCTAATAATATGCAACA 180
QY 121 CCAGGAAGGGGTTTGAAGTGGTGTGTCGCATAAAGAAAGTAAATATATGCAACA 180
DB 121 CCAGGAAGGGGTTTGAAGTGGTGTGTCGAATTAAGTATGATTTGAATCTAATAATATGCAACA 180
QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAACGATG 240
DB 181 CATTTATCGGAGTCTGTGAAGAGGGAGGTTCCACCATCTCAAGAGATGATTTCCAAAGTAGT 240

QY 241 CTCATCTGCAAAATGAACAACCTTGAACACTGAGGACACAGCCATGTATTACTGTGTGACA 300
DB 241 GTCTACTGCAAACTGAACAACCTTAGAGCTGAAGACACTGGCATTTATTACTGTGCGCA-- 298
QY 301 CGGGGGGCTTCAGGAGTTGACTATGCTATGGACTACTGGGGTCAAGAAACCTCACTACACC 360
DB 299 -GGGAGAGGTACTCTTATTACTATATACTATGGACTACTGGGGTCAAGAAACCTCAGTCA 357
QY 361 GTCTCTCTC 368
DB 358 GTATCTCTC 365

RESULT 10
AAA44338
ID AAA44338 standard; cDNA; 359 BP.
XX
AC AAA44338;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human secreted expressed sequence tag..SEQ ID NO:913.
XX
KW Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antinflammatory; cyostatic; antibacterial; antifungal;
KW antiviral; antidiabetic; antidiastmatic; vulnerary; antiparkinsonian;
KW antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW Parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200021991-A1.
XX
PD 20-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24206.
XX
PR 15-OCT-1998; 98US-0104436.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX
WPI: 2000-317938/27.
XX
Isolated polynucleotides, and encoded proteins, comprising secreted
expressed sequence tags (sESTs), useful for treating various disorders
such as autoimmune, infectious, and central nervous system disorders -
Claim 1; Page 437-438; 803pp; English.
XX
AAA43426 to AAA45925 represent specifically claimed secreted expressed
sequence tags (sESTs), isolated from human, mouse, chicken and rat
tissue sources. The sESTs can have a range of activities depending on
the tissues they were isolated from. The activities include:
chemokinetic; proliferative; immunomodulatory; haematopoietic;
chemokinetic; analgesic; haemostatic; thrombolytic; antinflammatory;
cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
antinflammatory; vulnerary; antitumor; osteopathic; neuroprotective;
nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
anticonvulsant; and antidepressant. The sESTs can be used for gene
therapy and in vaccines. The sESTs are useful as probes for the
```

CC identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AA045946 to AA445931 represent linker variants which are given  
CC in the exemplification of the present invention.

XX  
SQ Sequence 359 BP; 91 A; 66 C; 93 G; 109 T; 0 other;

Query Match 61.2%; Score 225.8; DB 21; Length 359;  
Best Local Similarity 95.1%; Pred. No. 8.9e-58;  
Matches 233; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATGGTGAGGCTAAAGGGTCAATGAAACTC 60  
DB 114 GAGTGCAGCTGTTGAGTCTGGTGGAGGATGGTGAGGCTAAAGGGTCAATGAAACTC 173  
QY 61 TCATGTGCAGCCTCTGGATTCACCTTCAATAACTACGCCATGAATGGGTCCGCCAGGCT 120  
DB 174 TCATGTGCAGCCTCTGGATTCACCTTCAATAACTACGCCATGAATGGGTCCGCCAGGCT 233  
QY 121 CCAGGAAAGGGTTGGAAATGGTGGTTCGCATAGAAAGTAAAGTAATAATATGCAACA 180  
DB 234 CCAGGAAAGGGTTGGAAATGGTGGTTCGCATAGAAAGTAAAGTAATAATATGCAACA 293  
QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240  
DB 294 TATTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 353  
QY 241 CTCTA 245  
DB 354 CTCGA 358

RESULT 11  
AAT79537  
ID AAT79537 standard; DNA: 5227 BP.

XX  
AC AAT79537;  
XX  
DT 06-MAR-1998 (first entry)  
XX  
DE Plasmid pTERMSC2H10myc3sCAM.  
XX  
KW Protein-protein interaction; interacting polypeptide;  
KW polyphage principle; plasmid pTERMSC2H10myc3sCAM; vector;  
KW Escherichia coli display; peptide library; ds.

XX Synthetic.

OS  
PN WC9732017-A1.

XX  
PD 04-SEP-1997.

XX  
PF 26-FEB-1997; 97WO-EP00931.

XX  
PR 26-FEB-1996; 96EP-0102852.

XX  
PA (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.

XX  
PI Ge L, Ilag V;

XX  
DR WPI; 1997-448687/41.

XX  
PT Identification of interacting polypeptide encoding nucleic acid  
PT sequences - e.g. to identify protein-protein interactions, which  
PT play an important role in biological processes

XX  
PS  
XX Example 8; Fig 25; 105pp; English.

CC Plasmid pTERMSC2H10myc3sCAM (AAT79537) can be used in a novel method  
CC for identifying nucleic acid sequences (NAS), which encode  
CC polypeptides capable of interacting with at least 1 of their fellow  
CC polypeptides. This comprises: (a) providing a 1st library of  
CC recombinant vectors containing genetically diverse NAS comprising a  
CC variety of NAS encoding polypeptides; (b) providing a 2nd library of  
CC recombinant vectors containing genetically diverse NAS comprising a  
CC variety of NAS encoding polypeptides capable of interacting with  
CC the polypeptides of (a), where the vectors employed for production  
CC of the recombinant vectors and/or recombinant inserts display  
CC properties that are phenotypically distinguishable from those used  
CC in (a), and where at least 1 of the properties displayed by the  
CC vectors and/or recombinant inserts used in steps (a) and (b), upon  
CC the interaction of a polypeptide from the 1st library with a  
CC polypeptide from the 2nd library, together generate a screenable or  
CC selectable property; (c) expressing members of the libraries of  
CC recombinant vectors or nucleic acid sequences of (a) and (b), in  
CC appropriate host cells so that at least 1 interaction is  
CC established; and (d) selecting for the generation of the  
CC screenable or selectable property representing the interaction of  
CC the polypeptides. The method can be used to identify protein-protein  
CC interactions.

XX  
SQ Sequence 5227 BP; 1232 A; 1265 C; 1355 G; 1375 T; 0 other;

Query Match 58.4%; Score 215.4; DB 18; Length 5227;  
Best Local Similarity 76.7%; Pred. No. 2.9e-54;  
Matches 283; Conservative 0; Mismatches 71; Indels 15; Gaps 1;

QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATGGTGAGGCTAAAGGGTCAATGAAACTC 60  
DB 1830 GAAGTTAAACTGGTGGAGTCTGGTGGAGGAGGCTGGTGAACCTGGAGGATCCATGAAACTC 1889  
QY 61 TCATGTGCAGCCTCTGGATTCACCTTCAATAACTACGCCATGAATGGGTCCGCCAGGCT 120  
DB 1890 TCCTGTGTTGCCCTCTGGAATCACTTTTCAGTAATACCGGATGAACCTGGGTCCGCCAGTCT 1949  
QY 121 CCAGGAAAGGGTTGGAAATGGTGGTTCGCATAGAAAGTAAAGTAATAATATGCAACA 180  
DB 1950 CCAGGAAAGGGGCTTGAGTGGTTCGTAATAGATTGAATCTAATAATATGCAACA 2009  
QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240  
DB 2010 CATTATGCCGAGTCTGTGAAGGGAGGTTCCACCATCTCAAGAGATGATTCACAAAGTAGT 2069  
QY 241 CTCTATCTGCAAACTTGAACAACTTGAAACTCAGGACACAGCCATCTATTACTGTGTGAGA 300  
DB 2070 GTCTACTCTGCAAACTGAACAACTTAAGAGCTGAAGACACTGSCATTTTATTACTG----- 2122  
QY 301 CGGGGGGCTTCAGGAGTGTGACTATGCTATGGACTACTGGGGTCCANGAACTCTACTCACC 360  
DB 2123 -----TAGAGGGGTTTCATATACTATAGACTACTGGGGTCCANGAACTCTAGTCACA 2174  
QY 361 GTCTCTCTCA 369  
DB 2175 GTCTCTCTCA 2183

RESULT 12  
AAS03515  
ID AAS03515 standard; cDNA: 369 BP.

XX  
AC AAS03515;

XX  
DT 29-AUG-2001 (first entry)

XX  
DE DNA encoding anti-adipocyte monoclonal antibody heavy chain, FAT 103.

XX  
KW Antibody; adipocyte; heavy chain; light chain; obesity; FAT;  
KW heart disease; complementarity determining region; CDR; ss.

XX OS Homo sapiens.  
XX PN WO200127279-A1.  
XX PD 19-APR-2001.  
XX PF P-PSDB; AAU02615.  
XX PR 11-OCT-2000; 2000WO-GB03900.  
XX PR 12-OCT-1999; 99US-0158812.  
XX PR (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
XX PI Edwards BM, Main SH, Vaughan TJ;  
XX WPI: 2001-282031/29.  
XX DR P-PSDB; AAU02615.  
XX Panel of specific binding members of antibody molecules which bind to  
PT whole adipocytes is used in the treatment of obesity and obesity  
PT related diseases -  
XX Disclosure; Page 164; 182pp; English.  
XX AAS03401-AAS03535 represent anti-adipocyte monoclonal antibody heavy  
CC and light chain coding sequences of the invention. The antibodies can be  
CC used in the treatment of obesity and obesity related diseases. The  
CC antibodies can be used to deliver drugs or pro-drugs directly to the fat  
CC mass of an obese patient or the antibody can be used as a therapeutic  
CC itself. Antibodies binding specifically to adipocytes can be used to  
CC activate the immune system to destroy the cells by complement mediated  
CC lysis. The antibodies may be labeled with a detectable label such as  
CC radiolabel, fluorescent or chemical group and used in methods of  
CC diagnosis in human subjects e.g. to determine the presence of adipocyte  
CC antigen on the surface of an adipocyte to detect or determine the  
CC presence or level of adipocytes in a cell or tissue sample. The  
CC antibodies can be used as an alternative means of treatment for obese  
CC patients other than undergoing surgery to remove excess fat. Antibodies  
CC for different types of fat deposits can also be produced e.g. intra-  
XX abdominal fat associated with heart disease.  
XX Sequence 369 BP; 81 A; 90 C; 122 G; 76 T; 0 other;  
SQ

Query Match 58.2%; Score 214.6; DB 22; Length 369;  
Best Local Similarity 74.2%; Pred. No. 2e-54;  
Matches 271; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTTGGTGGAGGCTAAAGGTCATTGAAACTC 60  
DB 1 GAGGTCCAGCTGGTGGAGTCTGGTGGAGGCTGGTCCAGGCTGGGGGTCCCTGAAACTC 60  
QY 61 TCATGTCGAGCTCTGGATTCACCTTCAATTAACGACGATGAGTGGTCCGCCAGGCT 120  
DB 61 TCTGTGAGCTCTGGTGGTTCACCTTCACTGAGTGGTCTGTGTGACATGGTCCGCCAGGCT 120  
QY 121 CCAGGAAGAGGTTTGGAAATGGGTTGCTCCGATAGAGAAAGTAAAGTAATATATCAACA 180  
DB 121 TCCGGGAAGGGCTGGAGTGGGTTGGCCGTATTAGAAGCAAGCTAACAGTTACCGGACA 180  
QY 181 TTTTATGCGGATTCAGTGAAGACAGGTTCCACCTCTCCAGAGATGATTCACAAAGCATG 240  
DB 181 GCATATGTCGCTCGGTGAAAGGCAAGTTCACCTCTCCAGAGATGATTCACAAAGCAAG 240  
QY 241 CTCATCTGCAATCAACAACTTGAAGAACTGAGGACAGGACGATCTATTACTGTGTGAGA 300  
DB 241 GCGTATCTGCAATGAACAGCTTGAAGAAACCGAGGACAGCGGCTGTATTACTATTACTAGA 300  
QY 301 CCGGGGGGTTACAGGATTTGACTATGCTATGGACTACTGGGGTCAAGGAACCTCACTACC 360  
DB 301 CTTGAGATAGCAGTGGCGGTATGGGGAGGAGTACTGGGGAGGCGGACCCCTGTGTACC 360  
QY 361 GTCTC 365  
|||||

Db 361 GTCTC 365  
RESULT 13  
ID ABK47271 standard; DNA; 761 BP.  
XX AC ABK47271;  
XX DT 18-JUN-2002 (first entry)  
XX Insulin/insulin-like growth factor receptor, DNA sequence #25.  
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin; receptor; gene therapy; diabetes;  
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke;  
KW diabetic neuropathy; gene; ss.  
XX Synthetic.  
XX WO200172771-A2.  
XX PD 04-OCT-2001.  
XX 29-MAR-2000; 2000WO-US08528.  
XX 29-MAR-2000; 2000WO-US08528.  
XX (DGIB-) DGI BIOTECHNOLOGIES LLC.  
XX (NOVO) NOVO NORDISK AS.  
XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;  
PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandecki WS;  
PI Hansen PH, Ravera M, Hsiao K;  
XX WPI: 2002-025774/03.  
XX Modulating insulin activity in mammalian cells, for treating e.g.  
PT diabetes and tumours, comprises using peptides that bind to insulin or  
PT insulin-like growth factor receptors  
XX Example 1; Figure 19; 390pp; English.  
XX The invention relates to a method of modulating insulin activity in  
CC mammalian cells by administering a peptide that binds the insulin  
CC receptor (IR). A composition containing a peptide, optionally expressed  
CC from gene therapy vectors, that binds to Site 1 of IR and an insulin  
CC agonist are useful for treating diabetes. Also, peptides that are  
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
CC receptor agonists are useful for treating neurological diseases,  
CC including stroke and diabetic neuropathy. The peptides are also useful in  
CC screening for compounds that bind to IR or IGF-1 receptor, potential  
CC therapeutics and research reagents. ABK47246-ABK47277 represent IR  
CC and/or IGF-1 receptor DNA sequences of the invention.  
XX Sequence 761 BP; 179 A; 201 C; 219 G; 162 T; 0 other;  
SQ

Query Match 58.2%; Score 214.6; DB 24; Length 761;  
Best Local Similarity 74.2%; Pred. No. 2.6e-54;  
Matches 271; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
QY 5 TCATCTGCTGGTGGAGTCTGGTGGAGGATTTGGTCCAGCCTAAAGGTCATTGAAACTCTCAT 64  
DB 23 TGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTAAAGCCTGGGGGTCCTTAGACTCTCTCT 82  
QY 65 GTGACGCTCTGGATTCACCTTCAATACTACGCCATGAATTTGGTTCGCCAGGCTCCAG 124  
DB 83 GTGACGCTCTGGATTCACCTTCAATACTACGCCATGAATTTGGTTCGCCAGGCTCCAG 142  
QY 125 GAAAGGGTTTGAATGGGTTGCTGTCATAGAAGTAAAGTAATAATATATGCAACATTTT 184



DR P-PSDB; AAU72865.

XX Multifunctional polypeptides comprising binding sites that specifically  
XX recognise extracellular groups of the NKG2D receptor complex and  
PT domains which function as receptors or ligands, useful for treating  
PT cancers and infectious diseases -

XX Example 7; Fig 16; 114pp: English.

XX The invention relates to a multifunctional polypeptide comprising a  
XX domain with a binding site that specifically recognises an extracellular  
CC group of the NKG2D receptor complex and a second domain which functions  
CC as a receptor or ligand. The polypeptide and its associated  
CC polynucleotide are used for the preparation of a pharmaceutical  
CC composition for the treatment of cancer, infections and/or autoimmune  
CC conditions. The cancer may be a tumour of the head and neck, stomach,  
CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,  
CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,  
CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.  
CC The infectious diseases can be caused by viruses, bacteria, fungi,  
CC protozoa or helminths. The autoimmune diseases include multiple  
CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior  
CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent  
CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and  
CC autoimmune hepatitis. The sequences represent DNA encoding the  
XX polypeptides of the invention.

XX Sequence 768 BP; 179 A; 192 C; 203 G; 194 T; 0 other;

Query Match 57.7%; Score 212.8; DB 24; Length 768;  
Best Local Similarity 76.4%; Pred. NO. 9e-54;  
Matches 278; Conservative 0; Mismatches 77; Indels 9; Gaps 1;

QY	6	GATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCTAAAGGGTCATTGAAACTCTCATG	65
DB	9	GCTGCTCGAGGAGTCTGGAGGAGGCTTGGTGCACCTGGAGGATCCATGAAACTCTCCTG	68
QY	66	TGCAGCCTCTGGATTCACTTCAATAACTACGCCCATGAATTGGTCCGCCAGCTCCAGG	125
DB	69	TGTTGCCCTCTGGATTCACTTTCAGTAACACTACTGGATGAACCTGGTCCGCCAGCTCCAGA	128
QY	126	AAAGGGTTTGGGAATGGGTTCCTCGCATAGAAGTAAAGTAATAATATGCAACATTTTA	185
DB	129	GAAGGGCTTGAATGGGTTCCTGAAATTAGATTGAAATCTAATAATTATGCAACACATTA	188
QY	186	TGCCGATTCAGTGAAGACAGCTTACCACATCTCCAGAGATGATTCACAAGCATGCTCTA	245
DB	189	TGCGGAGTCTGGAAGGGAGGTTCCACCATCTCAAGAGATGATTCAAAAGTAGTGTCTA	248
QY	246	TCTGCAATGAACAACTTGAAACTGAGGACACAGCCATGTATTACTGTGTGAGACGGGG	305
DB	249	CCTGCAATGAACAACTTAAGAGCTGAAGACACTGGCATTTATTACTGTACCAGGC----	304
QY	306	GGCTTCAGGGATTGACTATGCTATGAGTACTGGGTCAAGGAACCTCACTCACCGTCTC	365
DB	305	-----TCCCTACGGCTTGTCTATGACTACTGGGGCCCAAGGACCACCGTCAACCGTCTC	359
QY	366	CTCA	369
DB	360	CTCA	363

Search completed: November 27, 2002, 03:53:36  
Job time : 213.234 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query Match	Length	DB	ID	Description
1	284.4	77.1	865	13	BI150936	BI150936 602916745
2	230.4	62.4	932	12	BF135929	BF135929 601781259
3	229.8	62.3	700	12	BG757291	BG757291 602715376
4	220.2	59.7	1039	14	BQ937972	BQ937972 AGENCECOURT
5	218.4	59.2	522	10	AW401500	AW401500 UI-HF-BKO
6	212.8	57.7	552	17	BH319627	BH319627 CH230-119

RESULT 1	BI150936	865 bp	mRNA	linear	EST 05-JUL-2001			
LOCUS	602916745F1	NCI_CGAP_Lu29	Mus musculus	cDNA clone	IMAGE:5067346 5',			
DEFINITION	mRNA sequence.							
ACCESSION	BI150936							
VERSION	BI150936.1	GI:14610937						
KEYWORDS	EST.							
SOURCE	house mouse.							
ORGANISM	Mus musculus							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
REFERENCE	1 (bases 1 to 865)							
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .							
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)							
JOURNAL	Unpublished (1999)							
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs-f@mail.nih.gov">cgabbs-f@mail.nih.gov</a> ; Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM1181 row: 1 column: 11 High quality sequence stop: 797. Location/Qualifiers 1. 865							
FEATURES	source							









musculus cDNA clone F430102008 5', mRNA sequence.

BB843560  
 VERSION BB843560.1 GI:17081927  
 EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Akimura, T., Arahawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, K., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.,  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

JOURNAL  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 , M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.  
 e mouse tissues.

FEATURES  
 source  
 1. 359  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="F430102008"  
 /clone\_lib="RIKEN full-length enriched, 6 days neonate spleen"  
 /tissue\_type="spleen"  
 /dev\_stage="6 days neonate"  
 /note="vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dt)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000,

Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

Mol Genet 7: 1967-1978."

BASE COUNT 90 a 69 c 91 g 109 t  
 ORIGIN

Query Match 57.5% Score 212; DB 10; Length 359;  
 Best Local Similarity 93.6%; Pred. No. 5.3e-51;  
 Matches 221; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCCTAAAGGGTCAATTCATTGAACATC 60  
 II III III III III III III III III III III III III III III III III III III III  
 Db 124 GAGGTGCAGCTTGTGTGAGTCTGGTGGAGGATTGGTGCAGCCCTAAAGGATTCATTGAACATC 183  
 QY 61 TCATGTGCAGCTTCGGATTTCACCTTCAATTAAGTACCTAGGCATCAATTTGGTCCGCCAGGCT 120  
 III  
 Db 184 TCATGTGCCGCTCTGGTTTCACTTCAATACCTATGCATGCATCTGGTGGTCCGCCAGGCT 243  
 QY 121 CCAGGAAAAGGGTTTGAATGGTTCCTGCATAAGAAAGTAAAGTAATTAATTCACAACA 180  
 III  
 Db 244 CCAGGAAAAGGGTTTGAATGGTTCCTGCATAAGAAAGTAAAGTAATTAATTCACAACA 303  
 QY 181 TTTTATGCCGATTCAGTGAAGACAGAGTTTCACTTCATCTCCAGAGATTCATTCACAAAAG 236  
 I III III III III III III III III III III III III III III III III III III III  
 Db 304 TATTATGCCGATTCAGTGAAGACAGAGTTTCACTTCATCTCCAGAGATTCATTCACAAAAG 359

RESULT 8  
 BG759283  
 LOCUS 602710970r1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4851232 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG759283  
 VERSION BG759283.1 GI:14069936  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 711)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cchapbs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1693 row: 1 column: 17  
 High quality sequence stop: 711.  
 Location/Qualifiers  
 1. 711  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4851232"  
 /clone\_lib="NIH\_MGC\_48"  
 /tissue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: B-cells; Vector: pOT87; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

```

Average insert size 1.6 kb. Constructed by Life Technologies. Note: This is a NCI-CCAP Library."
BASE COUNT      152 a      128 c      133 g      147 t
ORIGIN

Query Match          54.8%; Score 202.2; DB 12; Length 560;
Best Local Similarity 75.68; Pred. No. 4.3e-48;
Matches 279; Conservative 0; Mismatches 83; Indels 7; Gaps 2;

QY   1 GAAGTGTGCTGGTGAGTCTGGTGGAGGATTGGTCGAGCCTAAAGGTCATTTGAAACTC 60
    ||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   117 GAAGTGAAGCTTGAGGAGTCTGGAGGAGG-TTGGTGCNACCTGGNGAGCATCAATAACTC 175

QY   61 TCATGTGCAGCCTCTGGATTCACTTCAATAAATCACGCCATGAATTCGGTCCGCCAGGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   176 TCCTGTGCTTGCCCTCGGATTCACPTTTTAGCAACTACTGGATGACCTCGGTCCGCCAGTCT 235

QY   121 CCAGGAAGAAGGTTTGGAAATGGGTTGCTCGCNTAGAGAGTAAGAGTANVAATATTGCAACA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   236 CCAGAGAAGAGGCGTGGATGGATTGCTGAAATTAANTTGCATCTGATTAATTTTGCACA 295

QY   181 TTTTATCCGATTCCAGTGAAGACAGAGGTTCACCATCTCCAGAGATGATTTACAAGCATG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   296 CATTTATCGGAGTCTGTGATAGGAGGTTCCACCATCTCAGAGAGATNTCCAAAAGTAGT 355

QY   241 CTCATCTGCAAAATGACAACCTTGAAAACTGAGGACACAGCCATCTATTACTGTGTGAGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   356 GTCTACCTGCAAAATGAACAACCTTAAGAGCTGAAGACACTGGCTTTTATTACTGT----- 409

QY   301 CGGGGGGCTTCAGGGATTGACTATGCTATGAGCTACTGGGGTCANGAACCTCACTCACC 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   410 ACGGGTCTACTATGGCCGAGGTATGCTATGGACTACTGGGGTCAANGAACCTCAGTCACC 469

QY   361 GTCCTCCTCA 369

```

```

RESULT 10
BG754897
LOCUS
DEFINITION
    BG754897
    861 bp mRNA linear EST 15-MAY-2001
    602714288F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4854764 5',
    mRNA sequence.
ACCESSION
    BG754897
VERSION
    BG754897.1 GI:14065550
KEYWORDS
    EST.
SOURCE
    human.
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 861)
    NTH-MGC http://mgc.nci.nih.gov/.
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgapbs-f@mail.nih.gov
    Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
    CDNA Library Preparation: Ling Hong/Rubin Laboratory
    CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
    DNA sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LNL at:
    http://image.llnl.gov
    Plate: LLCM1702 row: o column: 21
    High quality sequence stop: 854.
    Location/Qualifiers
        1..861
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4854764"
        /clone_lib="NIH_MGC_48"
        /tissue_type="primary B-cells from tonsils (cell line)"
        /lab_host="DH10B (phage-resistant)"
FEATURES
    source

```

```

/notes=Organ: B-cells; Vector: pORB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by EcoR- $\alpha$ T priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAC(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
255 c 231 g 177 t
255 c

```

BASE COUNT	198 a	255 c	231 g	177 t
ORIGIN				
Query Match	54.6%; Score 201.4; DB 12; Length 861;			
Best Local Similarity	73.7%; Pred. No. 8.6e-48;			
Matches 272; Conservative	0;	Mismatches 91;	Indels 6;	Gaps 1;

[illegible]

RESULT 11					
BQ711793					
LOCUS	BQ711793	966 bp	mRNA	linear	EST 16-JUL-2002
DEFINITION	AGENCOURT_8487710 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6300705 5', mRNA sequence.				

sequence.

ACCESSION B0711793  
VERSION B0711793.1 GI:21850692  
KEYWORDS EST.  
SOURCE human.

ORGANISM	<p>Homo sapiens  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p>
REFERENCE	<p>NIH-MGC <a href="http://imgc.nci.nih.gov/">http://imgc.nci.nih.gov/</a>.</p>
AUTHORS	<p>National Institutes of Health, Mammalian Gene Collection (MGC)</p>
TITLE	<p>Unpublished (1999)</p>
JOURNAL	<p>Contact: Robert Strausberg, Ph.D.</p>
COMMENT	<p>Email: <a href="mailto:cgapbs.rem@nih.gov">cgapbs.rem@nih.gov</a>  Tissue Procurement: Dr. Mark Watson  cDNA Library Preparation: Rubin Laboratory  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  DNA Sequencing by: Agencourt Bioscience Corporation  Clone Distribution: MGC clone distribution information can be  found through the I.M.A.G.E. Consortium/LLNL at:  <a href="http://image.llnl.gov">http://image.llnl.gov</a>  Plate: LLCN2515 row: g column: 10  High quality sequence spot: 561.  Location/Qualifiers</p>
FEATURES	

## FEATURES

source

1. .966  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6300705"  
/clone\_lib="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="CDNA from: spleen; Vector: pOH7; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and  
Superscript II RT (Life technologies). Note: this is a  
NIH\_MGC Library."  
291 c 260 g 202 t  
a

BASE COUNT	213 a	291 c	260 g	202 t
ORIGIN				

Query Match 54.5%; Score 201.2; DB 14; Length 966;  
Best Local Similarity 79.1%; Pred. NO. 1e-47;  
Matches 239; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy	1	GAAGTGA	TGCTGGT	GGAGTCTGGT	GGAGGATTCGT	GAGCCTAAACGGT	CATTGAAACTC	60
Db	120	GAAGTGAGT	GTGGAGTCCG	GGGAGGCTTGGT	CCAGCTGGGGGT	CCCTGAAACTC	179	
Qy	61	TCATGTG	CAGCCTCTG	GATTCACTT	CAATAACTAG	CCCATGAAT	TGGGTCCGCCAGGCT	120
Db	180	TCCTGTG	CAGCCTCTG	GGTTCACTT	CAGTGGCTT	ACTATG	CACATGGTCCGCCAGGCT	239
Qy	121	CCAGGAAG	GGTTTGG	NAATGGGTTGCT	CGCATAAAG	AGTAAAGTAA	TATATGCAACA	180
Db	240	TCCGGGA	AGAGGGTGG	AGTGGGTTGGCCG	TATTAGACG	CAAGCTCACAGT	TTACGGCGCA	299
Qy	181	TTTTTAT	CGCGATT	CAGTGAAG	ACAGGTTCCACCAT	CTCCAGAGAT	GATTTACAAAGCATG	240
Db	300	TCATATG	CTCGCTCG	GTGAAAGG	CAGGTTCCACCAT	CTCCAGAGAT	GATTTCAAAGAACACG	359
Qy	241	CTCTATC	TGCAAA	TGACAACTTGA	AAACTGAGG	ACACGCCCATG	TATTACTGTGTGAG	300
Db	360	GCGTATC	TGCAAA	TGACAAACCTG	AAACCGGAGG	ACGCCGCTG	TATTACTGTGTGTA	419
Qy	301	CG	302					
Db	420	GG	421					

RESULT 12				
BG684987				
LOCUS	BG684987	693 bp	mRNA	linear
DEFINITION	602636738F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4764440 5', mRNA sequence.			EST 01-MAY-2001

Accession  
Version  
Keywords  
Source

ORGANISM HOMO SAPIENS

ORGANISM: *Homo sapiens*  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 693)  
 REFERENCE: NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE: National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL: Unpublished (1999)  
 COMMENT: Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLC1M621 row: 1 column: 09  
High quality sequence stop: 691.

```
FEATURES
source
Location/Qualifiers
1..693
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:476440"
/clone_lib="NIH_MGC_48"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: B-cells; Vector: pONB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G)...Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."
BASE COUNT 151 a 191 c 203 g 148 t
ORIGIN
Query Match 54.4%; Score 200.8; DB 12; Length 693;
Best Local Similarity 74.6%; Pred. No. 1.2e-47;
Matches 282; Conservative 0; Mismatches 87; Indels 9; Gaps 2;
QY 1 GAAGTGATCTGGTGAGTCTGGTGAGGATTGGTGACGCCCTAAAGGGTCATTGAAACATC 60
DB 127 GAGGTGCAGCTGGTGAGTCTGGGGAGGCTTGGTAAAGCCTGGGGGTCCTTAGACATC 186
QY 61 TCATGTGCAGCCTCTGGATTACCTTCATTAACCTACGCCCATGAATGGGTCCGCCAGGCT 120
DB 187 TCTGTGCAGCCTCTGGATTACCTTCATTAACGCCCTGGATGAGCTGGGTCCGCCAGGCT 246
QY 121 CCAGGAAGGGTTGGATGGGTTGCTCGCATAAAGATAAAGTAAATATTATGCAACA 180
DB 247 CCAGGAAGGGCTGGAGTGGGTGGCCGCATCAACAGCAAAAGTATGATGGTGGCAACA 306
QY 181 TTTTATCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTACAAAGCATG 240
DB 307 GACTACGGTGCACCCGTGAAGGCAGATTCCACATCTCAAGAGATGATTTCAGAAACATG 366
QY 241 CTCATCTCAATGACAACTTGAACCTGAGGACAGCCATGTATTACTGT--GTCA 298
DB 367 CTGAATCTCAAAATGAACAGCCTGAAAGCGAGGACACAGCCGTGTATTATTGTACCA 426
QY 299 GACGGGGGCTTCAGGAGTTGACTATGCT-----ATGGACTACTGGGTCAAGGAACC 351
DB 427 GATAGCGGAGAGGAGGAGTACTTTCTCTCGGACTTTGACTACTGGGCCAGGGAACC 486
QY 352 TCACCTACCGCTCTCCTCA 369
DB 487 CTGGTCACGGTCTCCTCA 504
RESULT 13
BF119726 548 bp mRNA linear EST 24-OCT-2000
601758583F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3992829 5',
mRNA sequence.
DEFINITION
BF119726 BF119726 GI:10958675
ACCESSION
VERSION
KEYWORDS
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 548)
REFERENCE
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9208 row: e column: 22
High quality sequence stop: 539.
Location/Qualifiers
1..548
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3992829"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 146 a 116 c 138 g 148 t
ORIGIN
Query Match 54.0%; Score 199.4; DB 12; Length 548;
Best Local Similarity 78.4%; Pred. No. 2.8e-47;
Matches 239; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 1 GAAGTGATCTGGTGAGTCTGGTGAGGATTGGTGACGCCCTAAAGGGTCATTGAAACATC 60
DB 59 GAAGTGAAGTTTTCAGTAGTCTGGAGGAGGCTTGGTACAAACCTCGGAGGATCCATGAAACATC 118
QY 61 TCATGTGCAGCCTCTGGATTACCTTCATTAACCTACGCCCATGAATGGGTCCGCCAGGCT 120
DB 119 TCTGTGTTCCCTCTCGATTCACTTTCACTAATCTATCATGAACTGGGTCCGCCAGTCT 178
QY 121 CCAGGAAGGGTTTGGATGGGTTGCTCGCATAAAGATAAAGTAAATATTATGCAACA 180
DB 179 CCAGGAAGGGCTTCAGTGGGTGCTGGAATTAGATTGAAATCTAATAATTATGCAACA 238
QY 181 TTTTATCCGATTCAGTGAAGACAGGTTCCACATCTCCAGAGATGATTTCACAAAGCATG 240
DB 239 CATTATCGGAGTCTGTGAAGGGAGGTTCCACATCTCAAGAGATGATTTCACAAAGTACT 298
QY 241 CTCATCTCAATGAACAACTTGAACCTGAGGACAGCCATGTATTACTGTGTGAGA 300
DB 299 GTCTACCTCAATGAACAACTTAAGAGTTGAGACACTGGCATTTATTACTGTACCAG 358
QY 301 CGGGG 305
DB 359 CTGGG 363
RESULT 14
BQ706843 909 bp mRNA linear EST 16-JUL-2002
AGENCY: 8351705 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6282330
5', mRNA sequence.
DEFINITION
BQ706843 BQ706843 GI:21845742
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 909)
REFERENCE
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:35:53 ; Search time 41.3581 Seconds  
(without alignments)  
2736.194 Million cell updates/sec

Title: US-09-893-615-86

Perfect score: 369

Sequence: 1 GAAGTGTGCTGGTGAGTC.....CCTCACTACCGTCTCTCA 369

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300.4	81.4	357	1 US-08-442-542-1	Sequence 1, Appli
2	300.4	81.4	357	3 US-08-765-469-1	Sequence 1, Appli
3	300.4	81.4	1797	1 US-08-442-542-17	Sequence 17, Appl
4	300.4	81.4	1797	3 US-08-765-469-17	Sequence 17, Appl
5	288.6	78.2	345	3 US-08-767-128-35	Sequence 35, Appl
6	284.4	77.1	330	3 US-08-767-128-23	Sequence 23, Appl
7	227.2	61.6	366	3 US-08-483-749A-1	Sequence 1, Appli
8	213.6	57.9	372	3 US-08-767-128-37	Sequence 37, Appl
9	205	55.6	361	2 US-08-647-144-1	Sequence 1, Appli
10	204.2	55.3	738	2 US-08-224-591-13	Sequence 13, Appl
11	204.2	55.3	738	2 US-08-392-338A-22	Sequence 22, Appl
12	204.2	55.3	738	2 US-08-926-789-13	Sequence 13, Appl
13	204.2	55.3	738	3 US-09-166-750-22	Sequence 22, Appl
14	204.2	55.3	738	3 US-09-166-093-22	Sequence 22, Appl
15	204.2	55.3	738	3 US-09-172-019-22	Sequence 22, Appl
16	204.2	55.3	738	3 US-09-166-094-22	Sequence 22, Appl
17	204.2	55.3	738	5 PCT-US93-11138-13	Sequence 13, Appl
18	204.2	55.3	744	2 US-08-392-338A-12	Sequence 12, Appl
19	204.2	55.3	744	3 US-09-166-750-12	Sequence 12, Appl
20	204.2	55.3	744	3 US-09-166-093-12	Sequence 12, Appl
21	204.2	55.3	744	3 US-09-172-019-12	Sequence 12, Appl
22	204.2	55.3	744	3 US-09-166-094-12	Sequence 12, Appl
23	203	55.0	345	1 US-08-468-661-2	Sequence 2, Appli
24	203	55.0	345	1 US-08-466-272A-2	Sequence 2, Appli
25	203	55.0	345	1 US-08-478-857-2	Sequence 2, Appli
26	203	55.0	345	2 US-08-471-771-2	Sequence 2, Appli
27	203	55.0	345	3 US-09-130-783-2	Sequence 2, Appli

Sequence 21, Appl  
Sequence 20, Appl  
Sequence 14, Appl  
Sequence 14, Appl  
Sequence 14, Appl  
Sequence 14, Appl  
Sequence 14, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 59, Appl  
Sequence 33, Appl  
Sequence 43, Appl  
Sequence 3, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-442-542-1

; Sequence 1, Application US/08442542

; Patent No. 5686600

; GENERAL INFORMATION:

; APPLICANT: Carozzi, Nadine B.

; TITLE OF INVENTION: Antibodies which Bind to Insect Gut

; NUMBER OF SEQUENCES: 49

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ciba-gelgy Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NY

; COUNTRY: USA

; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/442,542

; FILING DATE: 16-MAY-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/267,641

; FILING DATE: 28-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Spruill, W. Murray

; REGISTRATION NUMBER: 32,943

; REFERENCE/DOCKET NUMBER: CGC 1750

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8615

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 357 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..357

; OTHER INFORMATION: /note= "3B1 heavy chain variable

; OTHER INFORMATION: region from pCIB4613"

; US-08-442-542-1



NAME: Sprull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CSC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1797 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:

NAME/KEY: CDS

LOCATION: 1..1797

OTHER INFORMATION: /note= "3B1 single chain antibody

OTHER INFORMATION: from PCIB4631"

US-08-442-542-17

Query Match 81.4%; Score 300.4; DB 1; Length 1797;  
Best Local Similarity 91.0%; Pred. No. 6.1e-89;  
Matches 335; Conservative 0; Mismatches 21; Indels 12; Gaps 1;

QY 2 AAGTCATGCTGCTGAGTCTGCTGGAGGATTGCTGCAGCCTAAAGGGTCATTGAAACTCT 61  
| | | | |  
Db 464 AGGTCAAACTGCAGAGTCTGCTGGAGGATTGCTGCAGCCTAAAGGGTCATTGAAACTCT 523  
| | | | |  
QY 62 CATGTGCAGCCCTCTGGATTCACTTCAATACTAGCCATGAATTGGTCCGCCAGGCTC 121  
| | | | |  
Db 524 CATGTGCAGCCCTCTGGATTCACTTCAATACTAGCCATGAATTGGTCCGCCAGGCTC 583  
| | | | |  
QY 122 CAGGAAGGGTTTGAATGGTTCCTGCATGAAGTAAGTAATTAATTATGCAACAT 181  
| | | | |  
Db 584 CAGGAAGGGTTTGAATGGTTCCTGCATGAAGTAAGTAATTAATTATGCAACAT 643  
| | | | |  
QY 182 TTTATGCCGATTTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTACACAAAGCATGC 241  
| | | | |  
Db 644 CTTATGGCGATTTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTACACAAAGCATGT 703  
| | | | |  
QY 242 TCTATCTGCAATGAACAACCTTGAATGAGGACAGGACAGCCATGATTACTGTGTGAGAC 301  
| | | | |  
Db 704 TCTATCTGCAATGAACAACCTTGAATGAGGACAGGACAGCCATGATTACTGTGTGAG-- 761  
| | | | |  
QY 302 GGGGGGCTTCAGGAGTTGACTATGCTATGGAGTCTGGGGTCAAGGAACCTCACTCACCG 361  
| | | | |  
Db 762 -----GGTAGTATACGGTGTCTATGGAGTCTGGGGTCAAGGAACCTCACTCACCG 811  
| | | | |  
QY 362 TCTCTCTCA 369  
| | | | |  
Db 812 TCTCTCTCA 819  
| | | | |

#### RESULT 4

US-08-765-469-17  
Sequence 17, Application US/08765469  
Patent No. 6069301

GENERAL INFORMATION:

APPLICANT: Carozzi, Nadine B.  
APPLICANT: Koziel, Michael G.  
TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
TITLE OF INVENTION: Proteins and their Use  
NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/087765.469  
FILING DATE:  
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/267,641

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sprull, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: CSC 1750

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8615

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 1797 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cdna

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1797

OTHER INFORMATION: /note= "3B1 single chain antibody

OTHER INFORMATION: from PCIB4631"

US-08-765-469-17

Query Match 81.4%; Score 300.4; DB 3; Length 1797;  
Best Local Similarity 91.0%; Pred. No. 6.1e-89;  
Matches 335; Conservative 0; Mismatches 21; Indels 12; Gaps 1;

QY 2 AAGTCATGCTGCTGAGTCTGCTGGAGGATTGCTGCAGCCTAAAGGGTCATTGAAACTCT 61  
| | | | |  
Db 464 AGGTCAAACTGCAGAGTCTGCTGGAGGATTGCTGCAGCCTAAAGGGTCATTGAAACTCT 523  
| | | | |  
QY 62 CATGTGCAGCCCTCTGGATTCACTTCAATACTAGCCATGAATTGGTCCGCCAGGCTC 121  
| | | | |  
Db 524 CATGTGCAGCCCTCTGGATTCACTTCAATACTAGCCATGAATTGGTCCGCCAGGCTC 583  
| | | | |  
QY 122 CAGGAAGGGTTTGAATGGTTCCTGCATGAAGTAAGTAATTAATTATGCAACAT 181  
| | | | |  
Db 584 CAGGAAGGGTTTGAATGGTTCCTGCATGAAGTAAGTAATTAATTATGCAACAT 643  
| | | | |  
QY 182 TTTATGCCGATTTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTACACAAAGCATGC 241  
| | | | |  
Db 644 CTTATGGCGATTTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTACACAAAGCATGT 703  
| | | | |  
QY 242 TCTATCTGCAATGAACAACCTTGAATGAGGACAGGACAGCCATGATTACTGTGTGAGAC 301  
| | | | |  
Db 704 TCTATCTGCAATGAACAACCTTGAATGAGGACAGGACAGCCATGATTACTGTGTGAG-- 761  
| | | | |  
QY 302 GGGGGGCTTCAGGAGTTGACTATGCTATGGAGTCTGGGGTCAAGGAACCTCACTCACCG 361  
| | | | |  
Db 762 -----GGTAGTATACGGTGTCTATGGAGTCTGGGGTCAAGGAACCTCACTCACCG 811  
| | | | |  
QY 362 TCTCTCTCA 369  
| | | | |  
Db 812 TCTCTCTCA 819  
| | | | |

#### RESULT 5

US-08-767-128-35  
Sequence 35, Application US/08767128  
Patent No. 6111079

GENERAL INFORMATION:

APPLICANT: WYLIE, DWANE E.

APPLICANT: LOPEZ, OSVALDO

APPLICANT: MURRAY, PETER JOSEPH

APPLICANT: GOEBEL, PETER

TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND

;; TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
;; NUMBER OF SEQUENCES: 46  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
;; STREET: 3100 No. 6111079west Center, 90 South Seventh St  
;; CITY: Minneapolis  
;; STATE: MN  
;; COUNTRY: USA  
;; ZIP: 55402  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSEQ Version 1.5  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/767,128  
;; FILING DATE:  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US96/09258  
;; FILING DATE: 05-JUN-1996  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/541,373  
;; FILING DATE: 10-OCT-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/462,798  
;; FILING DATE: 05-JUN-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Carter, Charles G.  
;; REGISTRATION NUMBER: 35,093  
;; REFERENCE/DOCKET NUMBER: 8648.49USF1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 612/371-5278  
;; TELEFAX: 612/332-9081  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 35:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 345 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Genomic DNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; FEATURE:  
;; NAME/KEY: Coding Sequence  
;; LOCATION: 1..345  
;; OTHER INFORMATION:  
;; US-08-767-128-35

Query Match 78.2%; Score 288.6; DB 3; Length 345;  
Best Local Similarity 89.7%; Pred. No. 2.2e-85;  
Matches 331; Conservative 0; Mismatches 14; Indels 24; Gaps 1;  
QY 1 GAAGTCATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCCTAAAGGGTCATTGAAATC 60  
Db 1 GAGGTGCAGCTGTTGAGTCTGGTGGAGGATTGGTGCAGCCCTAAAGGGTCATTGAAATC 60  
QY 61 TCATGTGCAGCTCTGGATTCACCTTCAATACTAGCCATGAATTGGTCCGCCAGGCT 120  
Db 61 TCATGTGCAGCTCTGGATTCACCTTCAATACTAGCCATGAATTGGTCCGCCAGGCT 120  
QY 121 CCAGGAAAGGGTTGGAATGGTGTGCTGCATAAAGTAAGTAATTAATTATGCAACA 180  
Db 121 CAAGGAAAGGGTTGGAATGGTGTGCTGCATAAAGTAAGTAATTAATTATGCAACA 180  
QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACATCTCCAGAGATGATTACAAAGCATG 240

Db 181 TATTATGCCGATTCAGTGAAGACAGGTTCCACATCTCCAGAGATGATTACAAAGCATG 240  
QY 241 CTCATCTCGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGTATTTACTGTGTGAGA 300  
Db 241 CTCATCTCGCAATGAACAACTTGAAGAACTGAGGACACAGCCATGTATTTACTGTGTGAGA 300  
QY 301 CGGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGGTCAAGGAACCTCACTCACC 360  
Db 301 CGG-----AGGGACTACTGGGGTCAAGGAACCTCACTCACC 336  
QY 361 GTCTCTCTCA 369  
Db 337 GTCTCTCTCA 345  
RESULT 6  
US-08-767-128-23  
; Sequence 23, Application US/08767128  
; Patent No. 6111079  
; GENERAL INFORMATION:  
; APPLICANT: WYLIE, DWANE E.  
; APPLICANT: LOPEZ, OSVALDO  
; APPLICANT: MURRAY, PETER JOSEPH  
; APPLICANT: GOEBEL, PETER  
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
; TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 6111079west Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/767,128  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09258  
; FILING DATE: 05-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/541,373  
; FILING DATE: 10-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/462,798  
; FILING DATE: 05-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carter, Charles G.  
; REGISTRATION NUMBER: 35,093  
; REFERENCE/DOCKET NUMBER: 8648.49USF1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/371-5278  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 330 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; HYPOTHETICAL: NO

; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 1..330  
; OTHER INFORMATION:  
US-08-767-128-23

Query Match 77.18; Score 284.4; DB 3; Length 330;  
Best Local Similarity 93.4%; Pred. No. 5.2e-84;  
Matches 297; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
Qy 16 GAGTCGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTCTCATGTGCAGCCTCT 75  
Db 1 GAGTCGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTCTCATGTGCAGCCTCT 60  
Qy 76 GGATTCACCTTCAATACTAGCCCATGAATTTGGGTCGCCAGGCTCCAGGAAGGGTTTG 135  
Db 61 GGATTCACCTTCAATACTAGCCCATGAATTTGGGTCGCCAGGCTCCAGGAAGGGTTTG 120  
Qy 136 GAATGGTGTCTCGCATAGAGTAAAGTAAATTAATTAATGCAACATTTTATGCCGATTCA 195  
Db 121 GAATGGTGTCTCGCATAGAGTAAAGTAAATTAATTAATGCAACATTTTATGCCGATTCA 180  
Qy 196 GTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATCTCTATCTGCAAAATG 255  
Db 181 GTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATCTCTATCTGCAAAATG 240  
Qy 256 ACAACTTGAAACTTGAGGACACACCCATGTATTACTGTGTGAGAGGGGGCTTCAGGG 315  
Db 241 ACAACTTGAAACTTGAGGACACACCCATGTATTACTGTGTGAGAGGGGGCTTCAGGG 300  
Qy 316 ATTGACTATGCTATGAC 333  
Db 301 GCCCAAGGACTCTGGTC 318

RESULT 7  
US-08-483-749A-1  
; Sequence 1, Application US/08483749A  
; Patent No. 6054561  
; GENERAL INFORMATION:  
; APPLICANT: RING, DAVID B.  
; TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY  
; TITLE OF INVENTION: MOLECULES SPECIFIC FOR CANCER ANTIGENS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CITRON CORPORATION  
; STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097  
; CITY: EMERYVILLE  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483.749A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAVERIDE, PAUL B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0508.008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2585  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 366 base pairs

; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..366  
US-08-483-749A-1

Query Match 61.6%; Score 227.2; DB 3; Length 366;  
Best Local Similarity 78.0%; Pred. No. 3.2e-65;  
Matches 287; Conservative 0; Mismatches 78; Indels 3; Gaps 1;  
Qy 1 GAAATGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTC 60  
Db 1 GAAATGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTC 60  
Qy 61 TCATGTGACGCTCTGGATTTCACCTTCAATAACTAGCCCATGAATTTGGGTCGCCAGGCT 120  
Db 61 TCATGTGACGCTCTGGATTTCACCTTCAATAACTAGCCCATGAATTTGGGTCGCCAGGCT 120  
Qy 121 CCAGGAAAGGGTTTGAATGGTCTGCATGAAGAAGTAAAGTAAATTAATTAATGCAACA 180  
Db 121 CCAGGAAAGGGTTTGAATGGTCTGCATGAAGAAGTAAAGTAAATTAATTAATGCAACA 180  
Qy 181 TTTTATGCCGATTTCAGTGAAGACAGAGTTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240  
Db 181 CATTATGCCGAGTCTGTGAAGGGAGGTTTCCACCATCTCAAGAGATGATTCACAAAGTAGT 240  
Qy 241 CTCATCTGCAANTGAACAACTTGAAGTCTGAGGACACAGCCATGTATTACTGTGTGAGA 300  
Db 241 GTCTACCTGCAACTGAACAACTTGAAGTCTGAGGACACAGCCATGTATTACTGTGTGAGA 298  
Qy 301 CGGGGGGCTTCAAGGATTGACTATGCTATGGGACTTACTGGGGTCAAGGAACCTCAGTCACC 360  
Db 299 -GGGAGAGGACTCTTACTATGACTATGACTATGACTATGACTATGACTATGACTATGACT 357  
Qy 361 GTCTCCTC 368  
Db 358 GTATCCTC 365

RESULT 8  
US-08-767-128-37  
; Sequence 37, Application US/08767128  
; Patent No. 6111079  
; GENERAL INFORMATION:  
; APPLICANT: WYLIE, DWANE E.  
; APPLICANT: LOPEZ, OSVALDO  
; APPLICANT: MURRAY, PETER JOSEPH  
; APPLICANT: GOEBEL, PETER  
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
; TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Walter & Schmidt  
; STREET: 3100 No. 6111079 West Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/767.128  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 424



Db 301 CAG-----GGAGGGGGCTATACCTTTGGAGCTATTTGGGGTCAAGGAACCTCAGTCACC 351

QY 361 GTCTCCTCA 369  
|||||

Db 352 GTCTCCTCA 360

RESULT 10  
US-08-224-591-13  
; Sequence 13, Application US/08224591  
; Patent No. 5856456  
; GENERAL INFORMATION:  
; APPLICANT: Whitlow, Marc  
; APPLICANT: Filpula, David  
; TITLE OF INVENTION: Linker For Linked Fusion Polypeptides  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/224,591  
; FILING DATE: Herewith  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/002,845  
; FILING DATE: 15-JAN-1993  
; APPLICATION NUMBER: US 07/980,529  
; FILING DATE: 20-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldstein, Jorge A.  
; REGISTRATION NUMBER: 29,021  
; REFERENCE/DOCKET NUMBER: 0977.1920002/JAG  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 738 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1..726)  
US-08-224-591-13

Query Match 55.3%; Score 204.2; DB 2; Length 738;  
Best Local Similarity 74.8%; Pred. No. 1.5e-57;  
Matches 276; Conservative 0; Mismatches 78; Indels 15; Gaps 1;

QY 1 GAAGTGATGCTGGTGAGTCTGGTGAGGAGTGGTGCAGCCCTAAAGGGTCATTGAAACTC 60  
|||||

Db 376 GAAGTAAACTGGATGAGACTGGAGGCGTGGTGCAACCTGGGAGGCCCATGAAACTC 435

QY 61 TCATGTGACGCTCTGGATTCACCTCAATAACTAGCCCATGAATGGGTCCGCCAGGCT 120  
|||||

Db 436 TCGTGTGTCCTCTGGATTCACCTTTAGTGACTACTGTGTAACCTGGGTCCGCCAGTCT 495  
|||||

QY 121 CCAGGAAGGGTTTGAATGGGTGCTCGCTAAGAAAGTAAATTAATATGCAACA 180  
|||||

Db 496 CCAGAAAGGAGCTGGAGTGGGTAGCCACAATAAGAACACCTTATTAATATGAAACA 555  
|||||

QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTTACCATCTCCAGAGATGATTCACAAAGCATG 240  
|||||

Db 556 TATTATTAGATTCTGTGAAGGCAGATTCAACATCTCAAGAGATGATTTCCAAAAGTACT 615

QY 241 CTCATCTGCAATGAACAACCTTGAACACTGAGGACACACACCATGTATTACTGTGTGAGA 300  
|||||

Db 616 GTCTACCTGCAATGAACAACCTTAAGAGTTGAAGACATGGGTATCTATTACTG----- 668

QY 301 CGGGGGGCTTCAGGAGTTGACTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACC 360  
|||||

Db 669 -----TACGGGTCTTACTATGATGGACTACTGGGGTCAAGGAACCTCAGTCACC 720

QY 361 GTCTCCTCA 369  
|||||

Db 721 GTCTCCTAA 729

RESULT 11  
US-08-392-338A-22  
; Sequence 22, Application US/08392338A  
; Patent No. 5869620  
; GENERAL INFORMATION:  
; APPLICANT: Whitlow, Marc  
; APPLICANT: Wood, James F.  
; APPLICANT: Hardman, Karl  
; APPLICANT: Bird, Robert  
; APPLICANT: Filpula, David  
; TITLE OF INVENTION: Multivalent Antigen-Binding Proteins  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/392,338A  
; FILING DATE: 22-FEB-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/989,846  
; FILING DATE: 20-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/796,936  
; FILING DATE: 25-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldstein, Jorge A.  
; REGISTRATION NUMBER: 29,021  
; REFERENCE/DOCKET NUMBER: 0977.0030007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 738 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..726  
US-08-392-338A-22

Query Match 55.3%; Score 204.2; DB 2; Length 738;  
Best Local Similarity 74.8%; Pred. No. 1.5e-57;  
Matches 276; Conservative 0; Mismatches 78; Indels 15; Gaps 1;

QY 1 GAAGTGATGCTGGTGAGTCTGGTGAGGAGTGGTGCAGCCCTAAAGGGTCATTGAAACTC 60  
|||||

```
Db 376 GAAGTTAACTGGATGAGACTGGAGGAGCTTGGTGCAACCTGGGAGGCCCATGAAACTC 435
Qy 61 TCATGTCAGCCTCGGATCACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 120
Db 436 TCCTGTGTGGCTCGGATTCACCTTTAGTGACTACTGGATGAACCTGGGTCCGCCAGTCT 495
Qy 121 CCAGGAAGGTTTGGATGGTTGCTCGCATAGAAGTAAAGTAATAATTTATGCAACA 180
Db 496 CCAGGAAGGACTGGAGTGGGTAGCACAAATTTAGAACAAACCTTATATTTATGAACA 555
Qy 181 TTTTATCCGATTCAGTGAAGACAGCTTCCACCTCTCCAGAGATGATTCACAAAGCATG 240
Db 556 TATTATTACAGATTCGTGAAAGGCAGATTCACCATCTCAAGAGATGATTCCAAAGAGTAGT 615
Qy 241 CTCTATCTGCAAACTTGAACACTTGAACACTGAGGACAGCATGTTACTGTGTGAGA 300
Db 616 GTCTACCTGCAATTAAGAGTTGAAGACATGGGTATCTATTTACTG----- 668
Qy 301 CGGGGGCTTCAGGATTTGACTATGCTATGCTATGCTAGGCTACCTCAAGAACTCTACTCACC 360
Db 669 -----TACGGGTTCTTACTATGTTAGTACTGCTAGTCTAGTCTAGTCAAGAACTCAGTCACC 720
Qy 361 GTCTCCTCA 369
Db 721 GTCTCCTAA 729

RESULT 12
US-08-926-789-13
; Sequence 13, Application US/08926789
; Patent No. 5990275
; GENERAL INFORMATION:
; APPLICANT: Whitlow, Marc
; APPLICANT: Filpula, David
; TITLE OF INVENTION: Linker For Linked Fusion Polypeptides
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,789
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/224,591
; FILING DATE:
; APPLICATION NUMBER: US 08/002,845
; FILING DATE: 15-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/980,529
; FILING DATE: 20-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0977.1920002/JAG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1..726)
US-08-926-789-13

Query Match 55.3%; Score 204.2; DB 2: Length 738;
Best Local Similarity 74.8%; Pred. No. 1.5e-57;
Matches 276; Conservative 0; Mismatches 78; Indels 15; Gaps 1;

Qy 1 GAAGTATGCTGGTGGAGTCTGGTGGAGGATTTGGTGCAGCCTAAAGSGTCAATTTGAAACTC 60
Db 376 GAAGTTAACTGGATGAGACTGGAGGAGCTTGGTGCAACCTGGGAGGCCCATGAAACTC 435
Qy 61 TCATGTCAGCCTCGGATTCACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 120
Db 436 TCCTGTGTGGCTCGGATTCACCTTTAGTGACTACTGGATGAACCTGGGTCCGCCAGTCT 495
Qy 121 CCAGGAAGGTTTGGATGGTTGCTCGCATAGAAGTAAAGTAATAATTTATGCAACA 180
Db 496 CCAGGAAGGACTGGAGTGGGTAGCACAAATTTAGAACAAACCTTATATTTATGAACA 555
Qy 181 TTTTATCCGATTCAGTGAAGACAGCTTCCACCTCTCCAGAGATGATTCACAAAGCATG 240
Db 556 TATTATTACAGATTCGTGAAAGGCAGATTCACCATCTCAAGAGATGATTCCAAAGAGTAGT 615
Qy 241 CTCTATCTGCAAACTTGAACACTTGAACACTGAGGACAGCATGTTACTGTGTGAGA 300
Db 616 GTCTACCTGCAATTAAGAGTTGAAGACATGGGTATCTATTTACTG----- 668
Qy 301 CGGGGGCTTCAGGATTTGACTATGCTATGCTATGCTAGGCTACCTCAAGAACTCTACTCACC 360
Db 669 -----TACGGGTTCTTACTATGTTAGTACTGCTAGTCTAGTCAAGAACTCAGTCACC 720
Qy 361 GTCTCCTCA 369
Db 721 GTCTCCTAA 729

RESULT 13
US-09-166-750-22
; Sequence 22, Application US/09166750
; Patent No. 6025165
; GENERAL INFORMATION:
; APPLICANT: Whitlow, Marc
; APPLICANT: Wood, James F.
; APPLICANT: Hardman, Karl
; APPLICANT: Bird, Robert
; APPLICANT: Filpula, David
; APPLICANT: Rollence, Michelle
; TITLE OF INVENTION: Methods for Producing Multivalent Antigen-Binding
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/166,750
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/392,338
; FILING DATE: 22-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/989,846
```



;; FILING DATE: 20-NOV-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/796,936  
;; FILING DATE: 25-NOV-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Goldstein, Jorge A.  
;; REGISTRATION NUMBER: 29,021  
;; REFERENCE/DOCKET NUMBER: 0977.003000C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 371-2600  
;; TELEFAX: (202) 371-2540  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 738 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..726  
US-09-166-750-22

Query Match 55.3%; Score 204.2; DB 3; Length 738;  
Best Local Similarity 74.8%; Pred. No. 1.5e-57;  
Matches 276; Conservative 0; Mismatches 78; Indels 15; Gaps 1;  
QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTC 60  
DB 376 GAAGTTAACTGGATGAGACTGGAGAGGCTTGGTGCACCTGGGAGGCCCATGAAACTC 435  
QY 61 TCATGTGAGCCTCTGGATTCACTTCAATAACTACGCCATGAATGGTCCGCCAGGCT 120  
DB 436 TCCTGTGTGCTCTGGATTCACTTTTAGTGACTACTGGATCACTGGTCCGCCAGTCT 495  
QY 121 CCAGAAAGGGTTTGGATGGTTCCTGCCATAAGAAAGTAAAGTAAATATGCAACA 180  
DB 496 CCAGAAAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 555  
QY 181 TTTTATGCGGATTCAGTGAAGACAGGTTTCACTCTCCAGAGATGATTTCACAAAGCATG 240  
DB 556 TATTATTAGATTCGTGTAAGGAGATTCACCTCTCAAGAGATGATTCCAAAGTAGT 615  
QY 241 CTCATCTGCAAAATGAACAACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 300  
DB 616 GTCTACCTGCAAAATGAACAACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 668  
QY 301 CGGGGGGCTTCAGGATTCAGTATGCTATGGACTACTGGGTTCAAGGAACCTCAGTCACC 360  
DB 669 -----TACGGGTCTTACTATGGTATGGACTACTGGGTTCAAGGAACCTCAGTCACC 720  
QY 361 GTCTCCTCA 369  
DB 721 GTCTCCTAA 729

RESULT 14  
US-09-166-093-22  
; Sequence 22, Application US/09166093  
; Patent No. 6027725  
; GENERAL INFORMATION:  
; APPLICANT: Whitlow, Marc  
; APPLICANT: Wood, James F.  
; APPLICANT: Hardman, Karl  
; APPLICANT: Bird, Robert  
; APPLICANT: Filpula, David  
; APPLICANT: Rollence, Michelle  
; TITLE OF INVENTION: Multivalent Antigen-Binding Proteins  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, NW  
; CITY: Washington  
; STATE: D.C.

;; COUNTRY: U.S.A.  
;; ZIP: 20005  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/166,093  
;; FILING DATE: Herewith  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/392,338  
;; FILING DATE: 22-FEB-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/989,846  
;; FILING DATE: 20-NOV-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/796,936  
;; FILING DATE: 25-NOV-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Goldstein, Jorge A.  
;; REGISTRATION NUMBER: 29,021  
;; REFERENCE/DOCKET NUMBER: 0977.003000B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 371-2600  
;; TELEFAX: (202) 371-2540  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 738 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; TOPOLOGY: both  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..726  
US-09-166-093-22

Query Match 55.3%; Score 204.2; DB 3; Length 738;  
Best Local Similarity 74.8%; Pred. No. 1.5e-57;  
Matches 276; Conservative 0; Mismatches 78; Indels 15; Gaps 1;  
QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACTC 60  
DB 376 GAAGTTAACTGGATGAGACTGGAGAGGCTTGGTGCACCTGGGAGGCCCATGAAACTC 435  
QY 61 TCATGTGAGCCTCTGGATTCACTTCAATAACTACGCCATGAATGGTCCGCCAGGCT 120  
DB 436 TCCTGTGTGCTCTGGATTCACTTTTAGTGACTACTGGATCACTGGTCCGCCAGTCT 495  
QY 121 CCAGAAAGGGTTTGGATGGTTCCTGCCATAAGAAAGTAAAGTAAATATGCAACA 180  
DB 496 CCAGAAAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 555  
QY 181 TTTTATGCGGATTCAGTGAAGACAGGTTTCACTCTCCAGAGATGATTTCACAAAGCATG 240  
DB 556 TATTATTAGATTCGTGTAAGGAGATTCACCTCTCAAGAGATGATTCCAAAGTAGT 615  
QY 241 CTCATCTGCAAAATGAACAACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 300  
DB 616 GTCTACCTGCAAAATGAACAACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 668  
QY 301 CGGGGGGCTTCAGGATTCAGTATGCTATGGACTACTGGGTTCAAGGAACCTCAGTCACC 360  
DB 669 -----TACGGGTCTTACTATGGTATGGACTACTGGGTTCAAGGAACCTCAGTCACC 720  
QY 361 GTCTCCTCA 369  
DB 721 GTCTCCTAA 729  
RESULT 15  
US-09-172-019-22

```

; Sequence 22, Application US/09172019
; Patent No. 6103889
; GENERAL INFORMATION:
; APPLICANT: Whitlow, Marc
; APPLICANT: Hardman, Karl
; APPLICANT: Bird, Robert
; APPLICANT: Filpula, David
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Single-Chain
; TITLE OF INVENTION: Antigen-Binding Proteins (As Amended)
; NUMBER OF INVENTIONS: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,019
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/392,338
; FILING DATE: 22-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/989,846
; FILING DATE: 20-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/796,936
; FILING DATE: 25-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0977.0030000D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both )
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..726
;
US-09-172-019-22

```

Query Match	55.3%	Score	204.2	DB	3	Length	738
Best Local Similarity	74.8%	Pred. No.	1.5e-57				
Matches	276	Conservative	0	Mismatches	78	Indels	15
Gaps	1						
Qy	1	GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTCGAGCCTAAAGGCGTCATTGAAACTC	60				
Db	376	GAAGTTAACTGGATGAGACTGGAGGAGGCTTGGTGGCACTGGGAGCCCATGAAACTC	435				
Qy	61	TCATGTCGAGCCTCTGGATTCACTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT	120				
Db	436	TCCTGTGTTGCTCTGGATTCACTTTTAGTGACTACTCTGGATGAACCTGGTCCGCCAGTCT	495				
Qy	121	CCAGGAAAGGTTTGGAAATGGTGTGCGATAAGAAGTAAACGCTAATAATTATTCGCAACA	180				
Db	496	CCAGAGAAAGGACTGGAGTGGGTAGCAACAATTAAGAACAACACCTTTATATTATGAACA	555				
Qy	181	TTTTATCCGATTCAGTGAAGACAGGCTTCACCATCTCCAGAGATGATTCACAAAGCATG	240				
Db	556	TATTATTTCAGATCTGTGAAAGGCAGATTCAACCATCTCAAGAGATGATTCACAAAGTAGT	615				

Qy	241	CTCTATCTGCAATGACAACTTGAACACTGAGGACACACCCATGCTATTACTCTGTGACA	300
Db	616	GTCTACTGCAATGACAACTTAAGATTGAAGACATGSGGTATCTTTACTG	668
Qy	301	CGGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGGTCAAGAACCTCACTCACC	360
Db	669	-----TACGGGTTCCTTACTATGGTATGGACTACTGGGGTCAAGAACCTCACTCACC	720
Qy	361	GTCTCCTCA	369
Db	721	GTCTCCTAA	729

Search completed: November 27, 2002, 03:55:13  
Job time : 49.8581 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 03:46:58 ; Search time 42.4323 seconds  
(without alignments)  
3349.042 Million cell updates/sec

Title: US-09-893-615-86  
Perfect score: 369  
Sequence: 1 GAAGTGATGCTGGTGAGTC.....CCTCACTCACCGTCTCCTCA 369

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 341543 seqs, 192557720 residues  
Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	100.0	369	10	US-09-893-615-86
2	279.4	75.7	466	10	US-09-881-823-15
3	269.2	73.0	443	10	US-09-840-459-96
c 4	269.2	73.0	443	10	US-09-840-459-99
5	249.2	67.5	357	10	US-09-840-459-97
c 6	249.2	67.5	357	10	US-09-840-459-103
7	201.4	54.6	420	10	US-09-286-240-3
8	198.2	53.7	1446	10	US-09-910-059-130
9	192.4	52.1	357	10	US-09-756-301A-4
10	192.4	52.1	357	10	US-09-927-703-4
11	192.4	52.1	357	10	US-09-766-535A-4
12	192.4	52.1	357	10	US-09-756-161A-4
13	192.4	52.1	357	12	US-10-010-229-4
14	192.4	52.1	357	12	US-10-043-450-4
15	192.4	52.1	357	12	US-10-044-534-4
16	191.4	51.9	424	10	US-09-730-857-62
17	190.4	51.6	1437	9	US-10-124-905-7
18	190	51.5	856	10	US-09-883-758-3
c 19	190	51.5	856	10	US-09-883-758-6

20	189.8	51.4	424	10	US-09-730-857-50	Sequence 50, Appl
21	189.8	51.4	424	10	US-09-730-857-64	Sequence 64, Appl
22	188.2	51.0	424	10	US-09-730-857-14	Sequence 44, Appl
23	188.2	51.0	424	10	US-09-730-857-48	Sequence 48, Appl
24	188.2	51.0	424	10	US-09-730-857-54	Sequence 54, Appl
25	188.2	51.0	424	10	US-09-730-857-58	Sequence 58, Appl
26	187	50.7	466	10	US-09-881-823-11	Sequence 11, Appl
27	186.6	50.6	424	10	US-09-730-857-28	Sequence 28, Appl
28	186.6	50.6	424	10	US-09-730-857-40	Sequence 40, Appl
29	184	49.9	530	10	US-09-844-684-14	Sequence 14, Appl
30	183.6	49.8	363	10	US-09-822-698A-4	Sequence 4, Appl
31	183.6	49.8	453	10	US-09-564-329A-14	Sequence 14, Appl
32	183.6	49.8	453	10	US-09-855-153-14	Sequence 14, Appl
33	183.6	49.8	453	10	US-09-854-811-14	Sequence 14, Appl
34	183.6	49.8	453	10	US-09-934-773-14	Sequence 14, Appl
35	183.6	49.8	453	10	US-09-963-620-14	Sequence 14, Appl
36	183.6	49.8	899	10	US-09-883-758-1	Sequence 1, Appl
c 37	183.6	49.8	899	10	US-09-883-758-5	Sequence 5, Appl
38	183.6	49.8	1143	10	US-09-822-698A-6	Sequence 6, Appl
39	183.6	49.8	1356	10	US-09-822-698A-27	Sequence 27, Appl
40	182.4	49.4	580	10	US-09-844-684-12	Sequence 12, Appl
41	182	49.3	406	9	US-10-046-935-2168	Sequence 2168, Ap
42	179	48.5	504	12	US-10-006-773-3	Sequence 3, Appl
43	178.6	48.4	423	7	US-08-779-784-3	Sequence 3, Appl
44	176.8	47.9	364	7	US-08-779-784-14	Sequence 14, Appl
45	176.4	47.8	532	9	US-09-187-693-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1  
US-09-893-615-86  
; Sequence 86, Application US/09893615  
; Patent No. US20020082395A1  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Gerald W.  
; Schuman, Richard F.  
; Wong, Hing  
; Stinson, Jeffrey L.  
; TITLE OF INVENTION: OPSONIC AND PROTECTIVE MONOCLONAL AND  
; CHIMERIC ANTIBODIES SPECIFIC FOR LIPOTEICHOIC ACID OF GR  
; POSITIVE BACTERIA  
; NUMBER OF SEQUENCES: 89  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &  
; DUNNER, LLP  
; STREET: 1300 I Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; Zip: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/893,615  
; FILING DATE: 29-Jun-2001  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elnaudi, Carol P.  
; REGISTRATION NUMBER: 32,220  
; REFERENCE/DOCKET NUMBER: 04995.0041-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 86:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 369 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

```
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..369
; SEQUENCE DESCRIPTION: SEQ ID NO: 86:
US-09-893-615-86

Query Match      100.0%; Score 369; DB 10; Length 369;
Best Local Similarity 100.0%; Pred. No. 2.9e-103;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTGTGCTGGTGGAGTCTGGTGGAGGATTGGTGAGGCTTAAAGGGTCATTGAAACTC 60
Db 1 GAAGTGTGCTGGTGGAGTCTGGTGGAGGATTGGTGAGGCTTAAAGGGTCATTGAAACTC 60
QY 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 120
Db 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 120
QY 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 120
Db 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 120
QY 121 CCAGGAAGGGTTTGGAAATGGTTCGCGATAAGAAAGTAAAGTAATAATATGCAACA 180
Db 121 CCAGGAAGGGTTTGGAAATGGTTCGCGATAAGAAAGTAAAGTAATAATATGCAACA 180
QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240
Db 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240
QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240
Db 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240
QY 241 CTCCTATCTGCAAAATGAACAACCTTCAAACTCAGGACACAGCCATGTATCTGTGTGAGA 300
Db 241 CTCCTATCTGCAAAATGAACAACCTTCAAACTCAGGACACAGCCATGTATCTGTGTGAGA 300
QY 301 CGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGGTCAAGGAACCTCACTCACC 360
Db 301 CGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGGTCAAGGAACCTCACTCACC 360
QY 361 GTCCTCTCA 369
Db 361 GTCCTCTCA 369

RESULT 2
US-09-881-823-15
; Sequence 15, Application US/09881823
; Patent No. US20020068066A1
; GENERAL INFORMATION:
; APPLICANT: SHI, WENYUAN
; APPLICANT: ANDERSON, MAXWELL
; APPLICANT: MORRISON, SHERIE
; APPLICANT: TRINH, RYAN
; APPLICANT: WIMS, LETITIA
; APPLICANT: CHEN, LI
; TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries
; FILE REFERENCE: 22851-032
; CURRENT APPLICATION NUMBER: US/09/881,823
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 07/378,577
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Murine
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14)..(442)
US-09-881-823-15

Query Match      75.7%; Score 279.4; DB 10; Length 466;
Best Local Similarity 86.7%; Pred. No. 6.8e-76;
Matches 320; Conservative 0; Mismatches 46; Indels 3; Gaps 1;

QY 1 GAAGTGTGCTGGTGGAGTCTGGTGGAGGATTGGTGAGGCTTAAAGGGTCATTGAAACTC 60
Db 1 GAAGTGTGCTGGTGGAGTCTGGTGGAGGATTGGTGAGGCTTAAAGGGTCATTGAAACTC 60
QY 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 120
Db 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 120
QY 121 CCAGGAAGGGTTTGGAAATGGTTCGCGATAAGAAAGTAAAGTAATAATATGCAACA 180
Db 121 CCAGGAAGGGTTTGGAAATGGTTCGCGATAAGAAAGTAAAGTAATAATATGCAACA 180
QY 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240
Db 181 TTTTATGCCGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240
QY 241 CTCCTATCTGCAAAATGAACAACCTTCAAACTCAGGACACAGCCATGTATCTGTGTGAGA 300
Db 241 CTCCTATCTGCAAAATGAACAACCTTCAAACTCAGGACACAGCCATGTATCTGTGTGAGA 300
QY 301 CGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGGTCAAGGAACCTCACTCACC 360
Db 301 CGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGGTCAAGGAACCTCACTCACC 360
QY 361 GTCCTCTCA 369
Db 361 GTCCTCTCA 369

RESULT 3
US-09-840-459-96
; Sequence 96, Application US/09840459
; Patent No. US20020150576A1
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 96
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-840-459-96

Query Match      73.0%; Score 269.2; DB 10; Length 443;
Best Local Similarity 94.0%; Pred. No. 8.5e-73;
Matches 280; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GAAGTGTGCTGGTGGAGTCTGGTGGAGGATTGGTGAGGCTTAAAGGGTCATTGAAACTC 60
Db 58 GAGGTGCAGCTTGTGAGTCTGGAGGAGGATTGGTGAGGCTTAAAGGGTCATTGAAACTC 117
QY 61 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 120
Db 118 TCATGTGCAGCCTCTGGATTACCTTCAATAAATACCTACGCCATGAATGGTCCGCCAGGCT 177
QY 121 CCAGGAAGGGTTTGGAAATGGTTCGCGATAAGAAAGTAAAGTAATAATATGCAACA 180
Db 121 CCAGGAAGGGTTTGGAAATGGTTCGCGATAAGAAAGTAAAGTAATAATATGCAACA 180
```

Db 178 CCAGGAAAGGTTTGGAAATGGTTGCTCGCATAAGAACTAAAAATAATAATTATGCAACA 237  
Qy 181 TTTTATGCCATTCAGTGAAGACAGAGTTTCCACCATCTCCAGAGATGATTCACAAGCATG 240  
Db 238 TATTATGCCATTCAGTGAAGACAGATACACCATCTCCAGAGATGATTCAGAAGATG 297  
Qy 241 CTCTATCTGCAATGAACAACTTGAACACTGAGGACACACCCATGATTTACTGTGTGA 298  
Db 298 CTCTTCTGCAATGAACAACTTGAACACTGAGGACACACCCATGATTTACTGTGTGA 355

## RESULT 4

US-09-840-459-99/c  
; Sequence 99, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: Larosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE REFERENCE: 1855.1052-012  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 99  
; LENGTH: 443  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-840-459-99

Query Match 73.0%; Score 269.2; DB 10; Length 443;  
Best Local Similarity 94.0%; Pred. No. 8.5e-73;  
Matches 280; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTTGGTGAGGCTTAAAGGTCATTGAAACTC 60  
Db 386 GAGTGCAGCTTGTGAGTCTGGAGGAGGATTTGGTGAGGCTTAAAGGTCATTGAAACTC 327  
Qy 61 TCATGTGACGCTCTGATTCACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 120  
Db 326 TCATGTGACGCTCTGATTCACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 267  
Qy 121 CCAGGAAAGGTTTGGAAATGGTTGCTCGCATAGAGTAAAGTAAATAATTATGCAACA 180  
Db 266 CCAGGAAAGGTTTGGAAATGGTTGCTCGCATAGAGTAAAGTAAATAATTATGCAACA 207  
Qy 181 TTTTATGCCATTCAGTGAAGACAGAGTTTCCACCATCTCCAGAGATGATTCACAAGCATG 240  
Db 206 TATTATGCCATTCAGTGAAGACAGATACACCATCTCCAGAGATGATTCAGAAGATG 147  
Qy 241 CTCTATCTGCAATGAACAACTTGAACACTGAGGACACACCCATGATTTACTGTGTGA 298  
Db 146 CTCTTCTGCAATGAACAACTTGAACACTGAGGACACACCCATGATTTACTGTGTGA 89

## RESULT 5

US-09-840-459-97  
; Sequence 97, Application US/09840459  
; Patent No. US20020150576A1

; GENERAL INFORMATION:  
; APPLICANT: Larosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 97  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanized heavy chain  
US-09-840-459-97

Query Match 67.5%; Score 249.2; DB 10; Length 357;  
Best Local Similarity 90.5%; Pred. No. 9.8e-67;  
Matches 266; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
Qy 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTTGGTGAGGCTTAAAGGTCATTGAAACTC 60  
Db 1 GAGGTGCAATTTGGTTGAGTCTGGAGGAGGATTTGGTGAAGCTTGGGGGCTCATTTGAGACTC 60  
Qy 61 TCATGTGACGCTCTGATTCACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 120  
Db 61 TCATGTGACGCTCTGATTCACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 120  
Qy 121 CCAGGAAAGGTTTGGAAATGGTTGCTCGCATAGAGTAAAGTAAATAATTATGCAACA 180  
Db 121 CCAGGAAAGGTTTGGAAATGGTTGCTCGCATAGAGTAAAGTAAATAATTATGCAACA 180  
Qy 181 TTTTATGCCATTCAGTGAAGACAGAGTTTCCACCATCTCCAGAGATGATTCACAAGCATG 240  
Db 181 TATTATGCCATTCAGTGAAGACAGATTCACCATCTCCAGAGATGATTCACAAGCATG 240  
Qy 241 CTCTATCTGCAATGAACAACTTGAACACTGAGGACACACCCATGATTTACTGT 294  
Db 241 CTCTATCTGCAATGAACAGCTTGAACAACTGAGGACACACCCGCTGATTTACTGT 294

## RESULT 6

US-09-840-459-103/c  
; Sequence 103, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:

; APPLICANT: Larosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02

```

: PRIOR APPLICATION NUMBER: 09/497,625
: PRIOR FILING DATE: 2000-02-03
: PRIOR APPLICATION NUMBER: 09/359,193
: PRIOR FILING DATE: 1999-07-22
: PRIOR APPLICATION NUMBER: 09/121,781
: PRIOR FILING DATE: 1998-07-23
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 103
: LENGTH: 357
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: - OTHER INFORMATION: Humanized heavy chain
US-09-840-459-103

```

[illegible]

## RESULT 8

```

US-09-910-059-130
; Sequence 130, Application US/09910059
; Patent NO. US20020142359A1
; GENERAL INFORMATION:
; APPLICANT: Copley, Clive G
; APPLICANT: Edge, Michael Derek
; APPLICANT: Emery, Stephen Charles
; TITLE OF INVENTION: Monoclonal Antibody to
; TITLE OF INVENTION: Their Therapeutic us
; FILE REFERENCE: 1991-209
; CURRENT APPLICATION NUMBER: US/09/910,059
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 09/171,945
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: PCT/GB97/01165
; PRIOR FILING DATE: 1997-04-29
; PRIOR APPLICATION NUMBER: GB 9703103.3
; PRIOR FILING DATE: 1997-02-14
; PRIOR APPLICATION NUMBER: GB9609405.7
; PRIOR FILING DATE: 1996-05-04
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 130
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: full-length human B7.
; NAME/KEY: CDS
; LOCATION: (16)..(1434)
; OTHER INFORMATION:
; US-09-910-059-130

```

Query Match 53.7% Score 198.2; DB 10; Length 1446;  
Best Local Similarity 73.2%; Pred. No. 6.2e-51;  
Matches 270; Conservative 0; Mismatches 93; Indels 6; Gaps 1;

Db 922 GAGTACAGTGCATCTGTGAAGGTCGGTTACCATTCTCCAGAGACAAAATCCCAAGCATC 981  
Oy 241 CTCTATCTGCAATCAACAACCTTGAAGAGTGGAGACAGCCATCTATTACTGTGTGAGA 300  
Db 982 CTCTATCTGCAATCAACAACCTTGAAGAGTGGAGACAGCCATCTATTACTGTGTGAGA 1041  
Oy 301 CGGGGGGCTTCAGGAGTTCATCTATGGAAGTGGGGTCAAGGAGAACTTCATCTCACC 360  
Db 1042 GATAGGGGCTACGGTCTTACT-----TTGACTACTGGGGCAAGGCACCACTCTCACA 1095  
Oy 361 GTCTCCTCA 369  
Db 1096 GTCTCCTCA 1104

RESULT 9

US-09-756-301A-4  
; Sequence 4, Application US/09756301A  
; Patent No. US20010027249A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Grayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-008  
; CURRENT APPLICATION NUMBER: US/09756.301A  
; CURRENT FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: U.S. 09/133,119  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: U.S. 08/570,674  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/324,799  
; PRIOR FILING DATE: 1994-10-18  
; PRIOR APPLICATION NUMBER: U.S. 08/192,102  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,861  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,093  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/010,406  
; PRIOR FILING DATE: 1993-01-29  
; PRIOR APPLICATION NUMBER: U.S. 08/013,413  
; PRIOR FILING DATE: 1993-02-02  
; PRIOR APPLICATION NUMBER: U.S. 07/943,852  
; PRIOR FILING DATE: 1992-09-11  
; PRIOR APPLICATION NUMBER: U.S. 07/853,606  
; PRIOR FILING DATE: 1992-03-18  
; PRIOR APPLICATION NUMBER: U.S. 07/670,827  
; PRIOR FILING DATE: 1991-03-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Mus Balb/c  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(357)

US-09-756-301A-4

Query Match 52.1%; Score 192.4; DB 10; Length 357;  
Best Local Similarity 72.7%; Pred. No. 2e-49;  
Matches 266; Conservative 0; Mismatches 91; Indels 9; Gaps 1;  
Oy 1 GAAGTGATCTGGTGGAGTCTGGTGGAGGATTTGGTGCAGCCCTAAAGGTCATTGAACTC 60  
Db 1 GAAGTGAAAGCTTGAGGAGTCTGGTGGAGGAGCTTGGTGCACCACTGGAGGATCCATGAAGCTC 60  
US-09-756-301A-4

Oy 61 TCATGTGCGAGCCTCGGATTCACCTTCAATAAATACGCCATGAATGGGTCCGCCAGGCT 120  
Db 61 TCCTGTGTGCTCTCGGATTCATTTTCAGTAACCACTGGATGAATGGGTCCGCCAGTCT 120  
Oy 121 CCAGGAAGAGGTTTGGATGGGTGCTGCCATGAAGAAGTAAAGTAATTAATTTATGCAACA 180  
Db 121 CCAGGAAGAGGCTTTCAGTGGGTGCTGAAATTAGATCAAAATCTATTAAATTTGCAACA 180  
Oy 181 TTTTATGCGGATTCAGTGAAGACAGAGTTTCCACCATCTCCAGAGATGATTTCCAAAAGTCATG 240  
Db 181 CATTATGCGGATCTGTGAAGGAGGTTTCACCATCTCAAGAGATGATTTCCAAAAGTGT 240  
Oy 241 CTCTATCTGCAAAATGAACAACCTTGAAGAGTGGAGACAGCCATCTATTACTGTGTGAGA 300  
Db 241 GTCTACCTGCAAAATGACCGACTTGAAGACTGAAGACACTGGCGTTTATTACTGTTCACG 300  
Oy 301 CGGGGGGCTTCAGGAGTTCATCTATGGAAGTGGGACTACTGGGGTCAAGGAACCTTCATCACC 360  
Db 301 AATTACTACGGTAGTACCTAC-----GACTACTGGGGCAAGGCACCACTCTCACA 351  
Oy 361 GTCTCC 366  
Db 352 GTCTCC 357

RESULT 10

US-09-927-703-4  
; Sequence 4, Application US/09927703  
; Patent No. US20020022720A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Grayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-013  
; CURRENT APPLICATION NUMBER: US/09/927,703  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: U.S. 09/756,398  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: U.S. 09/133,119  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: U.S. 08/570,674  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/324,799  
; PRIOR FILING DATE: 1994-10-18  
; PRIOR APPLICATION NUMBER: U.S. 08/192,102  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,861  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,093  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/010,406  
; PRIOR FILING DATE: 1993-01-29  
; PRIOR APPLICATION NUMBER: U.S. 08/013,413  
; PRIOR FILING DATE: 1993-02-02  
; PRIOR APPLICATION NUMBER: U.S. 07/943,852  
; PRIOR FILING DATE: 1992-09-11  
; PRIOR APPLICATION NUMBER: U.S. 07/853,606  
; PRIOR FILING DATE: 1992-03-18  
; PRIOR APPLICATION NUMBER: U.S. 07/670,827  
; PRIOR FILING DATE: 1991-03-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Mus Balb/c  
; FEATURE:  
; NAME/KEY: CDS





; PRIOR APPLICATION NUMBER: U.S. 08/010,406  
; PRIOR FILING DATE: 1993-01-29  
; PRIOR APPLICATION NUMBER: U.S. 08/013,413  
; PRIOR FILING DATE: 1993-02-02  
; PRIOR APPLICATION NUMBER: U.S. 07/943,852  
; PRIOR FILING DATE: 1992-09-11  
; PRIOR APPLICATION NUMBER: U.S. 07/853,606  
; PRIOR FILING DATE: 1992-03-18  
; PRIOR APPLICATION NUMBER: U.S. 07/670,827  
; PRIOR FILING DATE: 1991-03-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Mus Balb/c  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(357)  
US-09-756-161A-4

Query Match 52.1%; Score 192.4; DB 10; Length 357;  
Best Local Similarity 72.7%; Pred. No. 2e-49;  
Matches 266; Conservative 0; Mismatches 91; Indels 9; Gaps 1;

```
Qy 1 GAAGTGATGCTGGTGGAGCTCTGGTGGAGGATTGGTGCAGCCTAAAGGCTCATTTGAAACTC 60
    ||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GAAGTGAAGCTTGAGGAGCTCTGGAGGAGCTTGGTGCACCTGGAGGATCCATGAAACTC 60

Qy 61 TCATGTGCAGCCTCTGGATTACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TCCTGTGTGCTCTGGATTCAATTTTCAGTAACCACTGGATGGAAGTCCGCCAGTCT 120

Qy 121 CCAGAAAGGCTTGGATGGTTCGCGCATAGAACTAAGCTAAAGTAAATATGCAACA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CCAGAAAGGCTTGGATGGTTCGCGCATAGAACTAAGCTAAATATGCAACA 180

Qy 181 TTTATGCCGATTCAGTGAAGACAGGTTCCACCTCTCCAGAGATGATTCACAAAGCATG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CATTATGGGAGTCTGTGAAAGGGAGGTTCCACCTCTCAAGAGATGATTTCCAAAGTCT 240

Qy 241 CTCATCTGCAAACTGAACAACCTTGAAACTGAGGACAGGACCATGATTTACTGTGTGAGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GTCACCTGCAAACTGACCGACTTAGAAGTCAAGACACTGGCGTTTATTACTGTTCCAGG 300

Qy 301 CGGGGGCTTCAGGAGTTGACTATGCTATGAGTCTGGGTCAAGAACCTCACTCACC 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AATTACTAGGTAGTACCTAC-----GACTACTGGGGCCAAAGGACCACCTCTCACA 351

Qy 361 GTCCTCC 366
    |||||
Db 352 GTCCTCC 357
```

## RESULT 13

US-10-010-229-4  
; Sequence 4, Application US/10010229  
; Patent No. US20020114805A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Grayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-013  
; CURRENT APPLICATION NUMBER: US/10/010,229  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: US/09/927,703  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Mus Balb/c  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(357)  
US-10-010-229-4

Query Match 52.1%; Score 192.4; DB 12; Length 357;  
Best Local Similarity 72.7%; Pred. No. 2e-49;  
Matches 266; Conservative 0; Mismatches 91; Indels 9; Gaps 1;

```
Qy 1 GAAGTGATGCTGGTGGAGCTCTGGTGGAGGATTGGTGCAGCCTAAAGGCTCATTTGAAACTC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GAAGTGAAGCTTGAGGAGCTCTGGAGGAGCTTGGTGCACCTGGAGGATCCATGAAACTC 60

Qy 61 TCATGTGCAGCCTCTGGATTACCTTCAATAACTACGCCATGAATTTGGTCCGCCAGGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TCCTGTGTGCTCTGGATTCAATTTTCAGTAACCACTGGATGGAAGTCCGCCAGTCT 120

Qy 121 CCAGAAAGGCTTGGATGGTTCGCGCATAGAACTAAGCTAAAGTAAATATGCAACA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CCAGAAAGGCTTGGATGGTTCGCGCATAGAACTAAGCTAAATATGCAACA 180

Qy 181 TTTATGCCGATTCAGTGAAGACAGGTTCCACCTCTCCAGAGATGATTCACAAAGCATG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CATTATGGGAGTCTGTGAAAGGGAGGTTCCACCTCTCAAGAGATGATTTCCAAAGTCT 240

Qy 241 CTCATCTGCAAACTGAACAACCTTGAAACTGAGGACAGGACCATGATTTACTGTGTGAGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GTCACCTGCAAACTGACCGACTTAGAAGTCAAGACACTGGCGTTTATTACTGTTCCAGG 300

Qy 301 CGGGGGCTTCAGGAGTTGACTATGCTATGAGTCTGGGTCAAGAACCTCACTCACC 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AATTACTAGGTAGTACCTAC-----GACTACTGGGGCCAAAGGACCACCTCTCACA 351

Qy 361 GTCCTCC 366
    |||||
Db 352 GTCCTCC 357
```

## RESULT 14

US-10-043-450-4  
; Sequence 4, Application US/10043450  
; Patent No. US20020141996A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Grayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-013  
; CURRENT APPLICATION NUMBER: US/10/043,450  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: 09/927,703  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: U.S. 09/756,398  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: U.S. 09/133,119  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: U.S. 08/570,674  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/324,799  
; PRIOR FILING DATE: 1994-10-18  
; PRIOR APPLICATION NUMBER: U.S. 08/192,102  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,861  
; PRIOR FILING DATE: 1994-02-04

; PRIOR APPLICATION NUMBER: U.S. 08/192,093  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/010,406  
; PRIOR FILING DATE: 1993-01-29  
; PRIOR APPLICATION NUMBER: U.S. 08/013,413  
; PRIOR FILING DATE: 1993-02-02  
; PRIOR APPLICATION NUMBER: U.S. 07/943,852  
; PRIOR FILING DATE: 1992-09-11  
; PRIOR APPLICATION NUMBER: U.S. 07/853,606  
; PRIOR FILING DATE: 1992-03-18  
; PRIOR APPLICATION NUMBER: U.S. 07/670,827  
; PRIOR FILING DATE: 1991-03-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Mus Balb/c  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(357)  
US-10-043-450-4

Query Match 52.1%; Score 192.4; DB 12; Length 357;  
Best Local Similarity 72.7%; Pred. No. 2e-49;  
Matches 266; Conservative 0; Mismatches 91; Indels 9; Gaps 1;  
QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACATC 60  
DB 1 GAAGTGAACCTTGGAGGCTGGAGGAGGCTGGTGCACCTGGAGGATCCATGAACATC 60  
QY 61 TCATGTGACGCTCTGGATTCACCTTCAATAACTAGCCATGAATGGTCCGCCAGGCT 120  
DB 61 TCCTGTGTGCTCTGGATTCATTTTCAGTAAACCACTGGATGAACCTGGTCCGCCAGTCT 120  
QY 121 CCAGGAAGGGTTTGAATGGTGTCTCCGATGAAGAGTAAAGTAATTAATGCAACA 180  
DB 121 CCAGAGAAGGGGCTTGGAGTGGTGTCTGAATTAAGATCAAAATCTAATTAATCTGCAACA 180  
QY 181 TTTTATGCGGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240  
DB 181 CATTATGCGGAGTCTGTGAAGGGAGGTTCCACCATCTCAAGAGATGATTCCAAAGTGTCT 240  
QY 241 CTCATCTGCAAAATGAACAACTTGAACACTGAGGACACAGCCCATGATTTACTGTGTGAGA 300  
DB 241 GTCTACCTGCAAAATGACCACTTAAGAACTGAAGACACTGGCGTTTATTACTGTCCAGG 300  
QY 301 CGGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGTCAAGGAACCTCACTCACC 360  
DB 301 AATTACTAGGTAGTACCTAC-----GACTACTGGGGCCAGGCCACCTCTCTACA 351  
QY 361 GTCTCC 366  
DB 352 GTCTCC 357

RESULT 15  
US-10-044-534-4  
; Sequence 4, Application US/10044534  
; Patent No. US20020146419A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Grayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; FILE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-013  
; CURRENT APPLICATION NUMBER: US/10/044,534  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: 09/927,703

; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: U.S. 09/756,398  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: U.S. 09/133,119  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: U.S. 08/570,674  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/334,799  
; PRIOR FILING DATE: 1994-10-18  
; PRIOR APPLICATION NUMBER: U.S. 08/192,102  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,861  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,093  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: U.S. 08/010,406  
; PRIOR FILING DATE: 1993-01-29  
; PRIOR APPLICATION NUMBER: U.S. 08/013,413  
; PRIOR FILING DATE: 1993-02-02  
; PRIOR APPLICATION NUMBER: U.S. 07/943,852  
; PRIOR FILING DATE: 1992-09-11  
; PRIOR APPLICATION NUMBER: U.S. 07/853,606  
; PRIOR FILING DATE: 1992-03-18  
; PRIOR APPLICATION NUMBER: U.S. 07/670,827  
; PRIOR FILING DATE: 1991-03-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Mus Balb/c  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(357)  
US-10-044-534-4

Query Match 52.1%; Score 192.4; DB 12; Length 357;  
Best Local Similarity 72.7%; Pred. No. 2e-49;  
Matches 266; Conservative 0; Mismatches 91; Indels 9; Gaps 1;  
QY 1 GAAGTGATGCTGGTGGAGTCTGGTGGAGGATTGGTGCAGCCTAAAGGGTCATTGAAACATC 60  
DB 1 GAAGTGAACCTTGGAGGCTGGAGGAGGCTGGTGCACCTGGAGGATCCATGAACATC 60  
QY 61 TCATGTGACGCTCTGGATTCACCTTCAATAACTAGCCATGAATGGTCCGCCAGGCT 120  
DB 61 TCCTGTGTGCTCTGGATTCATTTTCAGTAAACCACTGGATGAACCTGGTCCGCCAGTCT 120  
QY 121 CCAGGAAGGGTTTGAATGGTGTCTCCGATGAAGAGTAAAGTAATTAATGCAACA 180  
DB 121 CCAGAGAAGGGGCTTGGAGTGGTGTCTGAATTAAGATCAAAATCTAATTAATCTGCAACA 180  
QY 181 TTTTATGCGGATTCAGTGAAGACAGGTTCCACCATCTCCAGAGATGATTCACAAAGCATG 240  
DB 181 CATTATGCGGAGTCTGTGAAGGGAGGTTCCACCATCTCAAGAGATGATTCCAAAGTGTCT 240  
QY 241 CTCATCTGCAAAATGAACAACTTGAACACTGAGGACACAGCCCATGATTTACTGTGTGAGA 300  
DB 241 GTCTACCTGCAAAATGACCACTTAAGAACTGAAGACACTGGCGTTTATTACTGTCCAGG 300  
QY 301 CGGGGGGCTTCAGGGATTGACTATGCTATGGACTACTGGGTCAAGGAACCTCACTCACC 360  
DB 301 AATTACTAGGTAGTACCTAC-----GACTACTGGGGCCAGGCCACCTCTCTACA 351  
QY 361 GTCTCC 366  
DB 352 GTCTCC 357

Search completed: November 27, 2002, 05:31:00  
Job time : 53.4323 secs